经济科学译库

策略博弈

(第二版)

Games of Strategy

(2nd Edition)

[美] 阿维纳什·迪克西特 Avinash Dixit
苏珊·斯克丝 Susan Skeath /著

蒲勇健 等/译

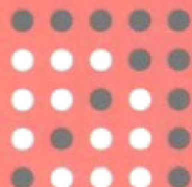
中国人民大学出版社

VISIT…

LANZAROTE Caliente.COM

Games of Strategy (2nd Edition)

著作权合同登记号
图字: 01-2004-6656号



迪克西特与斯克丝认识到了在大学课程计划的开始阶段教授博弈论概念的可能性：这是一个非常大的进步，值得赞许。通过使用《策略博弈》这本书，每一个地方的学生——正在安纳波利斯学习的年轻的军事战略家以及在芝加哥接受训练的经济理论家们——都可以对博弈论有一个初步的了解。大量的各种各样的解释性特例具有这样一种效果，即比起只有理论陈述而没有启发性例子来说，从该书里所学到的知识更加容易记住。

约翰·F·纳什
普林斯顿大学教授，诺贝尔经济学奖得主

了解博弈论将改变你整个一生中的思维方式。《策略博弈》为了解21世纪的新兴文化提供了一种令人赏心悦目的关键性构架。

保罗·A·萨缪尔森
麻省理工学院教授，诺贝尔经济学奖得主

教员们对本书第一版的褒奖

《策略博弈》是一部会为初学者们开辟一个崭新世界的详尽准确且令人感到非常愉悦的作品……每一个例子都非常艺术性地被挑选出来用以对策略性行为分析中的诸多难题之一加以解释……作者采用这样的思路和方法，使得其例子都是如此生动鲜明，从而产生了一种极为精细的和特别清晰的写作效果……这种非同寻常的力量组合使其成为一部理想的、内容自给的（self-contained）教材。

文森特·P·克劳福德
加州大学圣迭哥分校经济学系

该书是一部极棒的非技术性的博弈论导论性著作，行文令人赏心悦目且不失幽默。作者对该主题的处理是非数学化的，但仍不失严谨和准确性。该书作为一部非常全面的优秀教科书十分引人注目，并且它对于学习非技术性的入门水平的博弈论来说是一个非常好的选择。

加琳娜·A·施瓦茨
密歇根大学商学院教授

作为博弈论的一部导论性教材，该书绝无对手。

萨德什·穆疆达尔
南印第安纳大学商学院

我极为愿意赞美该作者写了一部启迪心智和有如此分析深度的，并且同时又是极具可读性和令人愉悦的书。干得好！

琳恩·佩帕尔
图辅（Tufts）大学经济学系

如果您想了解有关经济方面的图书信息，
请点击网址：www.crup.com.cn/jingji

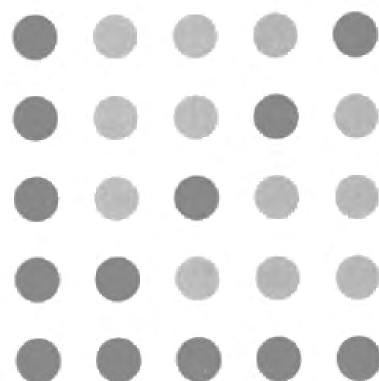


ISBN 978-7-300-10135-4/F·3459

定价: 65.00 元



经济科学译库



策略博弈（第二版）

[美] 阿维纳什·迪克西特 Avinash Dixit
苏珊·斯克丝 Susan Skeath /著

蒲勇健 等/译

Games of Strategy (2nd Edition)

中国人民大学出版社
·北京·

图书在版编目（CIP）数据

策略博弈：第2版/迪克西特，斯克丝著；蒲勇健等译.
北京：中国人民大学出版社，2009
（经济科学译库）
ISBN 978-7-300-10135-4

Ⅰ. 策…
Ⅱ. ①迪…②斯…③蒲…
Ⅲ. 对策论　高等学校—教材
Ⅳ. O225

中国版本图书馆CIP数据核字（2008）第202307号

经济科学译库
策略博弈（第二版）
［美］阿维纳什·迪克西特
　　苏珊·斯克丝　著
蒲勇健　等译

出版发行	中国人民大学出版社		
社　址	北京中关村大街31号	邮政编码	100080
电　话	010－62511242（总编室）		010－62511398（质管部）
	010－82501766（邮购部）		010－62514148（门市部）
	010－62515195（发行公司）		010－62515275（盗版举报）
网　址	http://www.crup.com.cn		
	http://www.ttrnet.com（人大教研网）		
经　销	新华书店		
印　刷	河北涿州星河印刷有限公司		
规　格	185mm×260mm　16开本	版　次	2009年1月第1版
印　张	33 插页 2	印　次	2009年1月第1次印刷
字　数	731 000	定　价	65.00元

版权所有　侵权必究　　印装差错　负责调换

译者序

我在大学时代就开始知晓博弈论，并对她产生了兴趣。但是，也只是到了20世纪90年代后期，才真正着手研究博弈论，同时开始给研究生和博士生开设博弈论课程。由于个人性格所致，一旦看上她就坠入爱河，愈陷愈深，不能自拔。于是，在研究上把主要的精力用在她身上，写文章，做课题，都逐渐向她靠拢。这么痴迷于博弈论，都是源于特别的个性。我本人天性比较内向，不善于与人打交道，于是在中学时代主攻数理化，尽管自己在人文方面也有着相当的兴趣。据说在美国有一句谚语说的是："如果你在哪方面不具天分的话，你就会成为那方面的专家（One would be an expert in the field in which he has less endowment)。"我在那时就突发奇想，认为有一天由于科学的发展，人们也能够像数理化那样运用数学方法去研究人类行为。而倘若真的能够这样，我就会去从事这样的研究。这样的想法在当时只是一种不着边际的空想，没想到接触博弈论之后，发现她正是干这种事情的。你想想，在这样的情况下我能不被她迷住吗！

除了对人类行为进行严格的数理分析外，博弈论已经成为了"经济学帝国主义"的急先锋。她入侵法律、政治科学、社会学、历史、语言等其他社会科学领域，甚至在诸如生物学、人工智能和火箭工程等自然科学及技术领域也大展雄姿。在大学时代曾料想总有那么一天会有一种研究自然科学与社会科学的统一方法论，现在看起来非博弈论莫属了，尽管她离这个目标实际上还远得很。

在大学时代，刚刚知道博弈论时就把她与打牌、下棋和战争角力联系在一起，觉得十分好玩。后来随着对博弈论研究的不断深入，对她的了解和熟悉程度

也不断提高，感觉到博弈论除了纯粹的好玩之外，其内涵也十分深邃厚重，散发着令人沉醉的魅力。她是如此美妙，令我不知不觉已拜倒在她的石榴裙下了！

在经历了数学和物理学学习研究之后，最终情归博弈论，从此此心专一，新鲜如初。七八年下来，在博弈论方面的著述也有了一些积累，在国内外也结交了许多致力于博弈论研究的学者，成为了博弈论一族中的一员。

2005 年，中国人民大学出版社的马学亮先生来电话，说他们出版社买了美国普林斯顿大学阿维纳什·迪克西特（Avinash Dixit）与苏姗·斯克丝（Susan Skeath）的《策略博弈》（*Games of Strategy*）一书第二版的版权，并告知我中国人民大学出版社决定由我来主持该书中文版的翻译。听到这个消息，我顿时兴奋不已。因为说起该书作者阿维纳什·迪克西特，其大名早已如雷贯耳。在博弈论及新经济增长理论的文献中（本人在转人博弈论领域之前曾专注于新经济增长理论领域），此君大名早已耳熟能详。2004 年，本人在香港大学工作时曾在香港大学校内书店看到该书的第一版，从其作者前言中知第二版的由来如下：阿维纳什·迪克西特曾经在 20 世纪 80 年代与巴里·奈尔伯夫（Barry Nalebuff）合作写过一本《策略思维》（*Thinking Strategically*），该书是 20 世纪 80 年代国际上十分有影响的供 MBA 学生和广大博弈论爱好者阅读的博弈论通俗教材。该书由中国人民大学出版社翻译出版了中文本，在国内也有相当大的影响。在更早的时候，也就是 1999 年，我在香港中文大学工作时曾在香港博弈论学者赵耀华（Steve Chiu）博士那里借过这本书的台湾译本，台湾的译名叫《大谋略——游戏理论的全方位运用》。但是，正如作者在《策略博弈》（第一版）前言里所说，该书有一个缺点，即在没有给出有关博弈论基本框架及概念的情况下讲故事，使得许多缺乏博弈论基础知识的读者在阅读该书时感到别扭。也就是在这样的市场信息反馈下，迪克西特又与斯克丝合作撰写了《策略博弈》这本书。《策略博弈》在其内容体系上基本克服了《策略思维》的上述不足，在有关博弈论的基本概念介绍上十分系统和通俗易懂。现在我们翻译的这本书是《策略博弈》的第二版，与其第一版相比，正如作者在前言中所言，最为实质性的变化出现在由第 3 章到第 8 章（在第一版里是第 3 章到第 7 章）构成的第二部分里。作者在第二版中重新撰写了这些章节，几乎是完整地详细描述了序贯和同时行动博弈的基本理论。作者们重新对素材的顺序加以了安排，以增进流畅性，删除或更换了某些太难或者是直接相关性稍弱的章节，并且还对作者借以提出理论性概念的大多数例子进行了更改或改进。

无论是《策略思维》还是《策略博弈》，都是以轻松的故事和颇为幽默的文笔为读者们讲授博弈论的精彩内容。按照作者自己的说法，他们撰写这本书的目的是为了使得这样一件事情成为可能，即不需要经济学或政治科学或生物学知识的先期准备，而只需要最低程度的中学数学知识，就可以为一年级或者是二年级的大学生讲授导论或“原理”水平的博弈论课程。事实上，目前绝大多数博弈论教材都基本上是为研究生准备的，真正可供本科生阅读的博弈论教科书还是凤毛麟角的。《策略博弈》是目前国际上较为流行的既适于 MBA 学生阅读，又可作为本科生博弈论教材的为数不多的佳品。

迪克西特教授并不仅仅是一位优秀的经济学或者博弈论科普作家，他的著述

也不仅限于经济学或博弈论教科书。他是当代著名经济学家，普林斯顿大学经济系终身教授。他1944年出生于印度孟买，早年研究数学和运筹学，曾先后在孟买大学和剑桥大学获得理学士（1963）和文学士（1965）学位。与经济学大师弗兰克·费希尔（Frank Fisher）的一次交谈，使迪克西特意识到自己在经济学研究中的潜力，于是毅然放弃了数学领域的奋斗，转向经济学。1968年，他在美国麻省理工学院（MIT）获得经济学博士学位。从那以后，迪克西特在经济学的广泛领域开始了漫长的征程，在微观经济理论、博弈论、国际贸易、产业组织、增长与发展理论、公共经济学、政治经济学以及新制度经济学等诸多领域都有杰出的表现。1996年他还被挪威经济与工商管理学院授予经济学荣誉博士学位，2001年担任计量经济学会会长，2002年担任美国经济学联合会副会长。他曾于1992年当选为美国艺术和科学研究院院士，2005年当选为国家科学院院士。

迪克西特著述颇丰，发表有超过150篇的期刊论文，其中有相当部分论文见于AER、JPE、QJE、*Econometrica*等顶级期刊。他也曾长期担任*Econometrica*、*Journal of Economic Theory*等杂志的副主编。漫长的教学生涯中，他也写下了许多著作和教材，包括《国际贸易理论》（与维克多·诺尔曼合作）、《经济理论中的最优化方法》、《策略思维》（与奈尔伯夫合作）、《不确定条件下的投资》（与罗伯特·平狄克合作）、《经济政策的制定：交易成本政治学的视角》、《策略博弈》（与苏姗·斯克丝合作），以及最近的《法律缺失与经济学：可供选择的经济治理方式》。中国读者对他的著作已不陌生，因为其中的大部分已经有中文版。

在近四十年的经济学研究生涯中，迪克西特始终活跃在学术研究的前沿，可谓经济学界的常青树。这可能与迪克西特独特的研究风格和方法有关。他曾对青年学者的研究提出这样的建议：不要太在意课题的社会意义有多大，要真正关注的是这个课题能否充分发挥你的聪明才智和丰富的、创造性的想象力。确定某个课题是否适合自己的一个重要标志，是看它能否使自己兴奋。这确是他的经验之谈，因为早年的迪克西特也研究了许多流行于其时代的问题，但他真正具有代表性的成果比如垄断竞争下产品差异理论、寡头垄断进入壁垒、国际贸易理论创新以及不可逆投资理论，都并不是时代潮流的产物。事实上，经济学中许多在后来被证明为最有价值的重要理论比如理性预期、信息与激励经济学、博弈论等，与其所产生时代的时政和潮流并无多大联系。所以，许多年以后，迪克西特这样总结说："追求潮流、追求社会的热点往往不是经济学理论研究的方向。"

迪克西特甚至不主张频繁地参加各种研讨会。他这样告诫年轻的学者："不要热衷于参加研讨会，除非他们指定的题目刚好是你所选的课题。因为写研讨会的论文会占用你太多的时间，而你本来可以利用这段时间做些你更感兴趣的工作。坚持你的工作，如果成功的话，几年以后人们将会就你的题目举行研讨会。同时，你在自己专长领域的成功从物质上的回报讲，也会超过参加十个研讨会带来的荣誉。"

国内的学者，一些人太过浮躁，急于发表文章，而另一些则太过完美主义，总是希望将文章修改到尽善尽美以致长期无文章可投。在迪克西特看来，这两种态度都是不可取的。他认为一个学者应该对他的文章进行必要的修改，以提高质量，对于已决定长期以学术为职业的人来说，尤其应注意提高文章的质量，这样

才有利于在圈内树立良好的声誉。但是，完美主义也并不可取，修改文章也应考虑到其机会成本。当你的文章修改到一定水平后，凭你个人的努力要想再提高则要花费更多的时间，而这段时间你本可以做其他更有效率的工作。追求一篇文章尽善尽美对于大多数人来说都是不切实际的。从个人角度来看，它可能会延缓个人的研究工作进程，增加整个研究体系的风险；从社会角度来说，某些并不完美的东西发表之后，其他人可能从中发现其比较优势，从而完成进一步的改进。

最后一个问题是学者们都感兴趣的：学者是否应该宣传自己工作的重要性？迪克西特本人非常谦虚，也从不过分地宣传自己的工作。但是，他认为适度的宣传是很重要的。一方面，如果你做了重要工作而无人知晓，那么你的成果就不能起到推进学术发展的作用；倘若大家都很保守低调，捂紧自己的成果保持缄默，经济学的发展就会停滞不前。但是，如果过分宣传自己工作的重要性，则可能使一个人在圈子内的声誉变得很糟糕，从而影响到将来的工作被认可的可能性。究竟如何宣传自己的工作？作为博弈论专家的迪克西特给出的最优策略是：宣传时要把重要性讲得比实际的重要性稍微小一点；如果非要夸大一点的话，也要采用比较间接的聪明办法。

即使如迪克西特这样的大家，也十分注意把经济学或博弈论看成是一门快乐的科学，这是十分难能可贵的。

据说英国近代哲学家边沁认为人们的行为是由最大化其快乐的原则来支配的。这样一种理论后来又成为现代经济学最基本的假定，也就是“效用最大化假定”。是的，至少按照我自己的想法来说，寻求快乐是我们进行决策的出发点，只不过有些是直接地实现快乐，而另外一些是间接地实现快乐而已。因此，在学习方面，也应该按照快乐的原则来选择学习的方式方法。快乐地学习，是我自少年时代起就身体力行的。在我的少年时光里，由于性格内向，不太合群，就在书海里寻找快乐。于是，快乐学习就无师自通。现在，我将这种原则已经扩展到更加广泛的“快乐地学习，快乐地工作，快乐地生活”，朋友们将其概括为“3H”原则（Learning Happily，Working Happily，Live Happily）。我本人曾经读过三个学位，三个专业，在不同的专业里获得了不同的学位。大学学数学，硕士研究生读的是物理，而博士研究生学经济管理，都有一个共同的原因，就是快乐原则。在不同的时期产生了不同的兴趣，于是从学习不同领域里的知识中获得了不同的快乐。

博弈论是数学家发明的，而数学又往往给人以艰深晦涩难懂的感觉，好像与快乐是沾不上边的。其实不然，博弈论是非常有趣、非常好玩的。博弈论的英文名称是“game theory”，一般认为在这里不能把 game 理解为“游戏”，但我却认为不妨把 game 就看成是游戏，博弈就是玩游戏，所以 game theory 也就是玩游戏的理论，因而也就是一种制造快乐的理论！数学家其实在数学研究中是十分快乐的，这一点本人在学习和研究数学的过程中深深体会过。只不过数学家是用数学语言表达出博弈论的快乐，他们就像高雅音乐的爱好者那样从欣赏交响乐中获得快乐。同样，博弈论也完全可以不使用复杂的数学来加以表述，这就是讲故事，就像我们不一定要欣赏交响乐，从通俗音乐中也可感受到音乐的美并由此带来快乐一样。不仅如此，其实科学本身在一定程度上就是从寻求快乐中产生的。

这么说有何依据呢？我们知道，人类科学的雏形基本上是在古代希腊形成的（中国古代在技术上有很大贡献，但在纯粹科学上的贡献则要让位于古希腊），而古希腊人生性乐观，豪放快乐，是他们最早在纯粹科学的探索上为后世确立了基本的规范。我们来看看法国哲学家丹纳在其名著《艺术哲学》中对古希腊人的描述吧："……腓尼基人长于经商，有一套数学用来算账。埃及人会丈量，凿石头，有一套几何学，在尼罗河一年一度的洪水之后用来恢复田地的疆界。希腊人向他们学了这些技术和方法，还嫌不够。他们不能满足于工商业上的应用；他们生性好奇，喜欢思索；他们要知道事物的原因和理由，他们追求抽象的证据，探索从一个定理发展到另一个定理的观念有哪些微妙的阶段。基督降生前六百多年，泰勒斯已在论证三角形两等边对角相等。据古人传说，毕达哥拉斯发现了'从直角三角形之弦引伸的方形，等于其他两边引伸的两个方形之和'的定理，欣喜若狂，甚至许下愿心要大祭神明。他们感兴趣的是纯粹的真理。柏拉图看到西西里的数学家把他们的发现应用于机器，责备他们损害了科学的尊严；按照他的意思，科学当以研究抽象的东西为限。的确，希腊人不断推进科学，从不考虑实用。他们对圆锥曲线的特性的研究，直到1700年后开普勒探求行星运动的规律时，才得到应用。几何学是我们一切正确的科学的基础，他们在这方面的正确分析，使英国至今还用欧几里得几何作为学校教本。分析各种观念，注意它们的隶属关系，建立它们的连锁，不让其中缺少一个环节，使整个连锁有一项颠扑不破的定理或是大家熟悉的一组经验做根据，津津有味地铸成所有的环节，把它们结合起来，加总，进行检验，唯一的动机是使这些环节越多越好，越紧密越好。这便是希腊人在智力方面的特长。他们为思想而思想，为思想而创造科学。我们今天建立的科学没有一门不建立在他们所奠定的基础之上；第一层楼往往是他们盖造的，有时甚至整整的一间屋子。发明家前后踵接！数学从毕达哥拉斯到阿基米得，天文学方面从泰勒斯与毕达哥拉斯到喜帕恰斯与托勒密；自然科学从希波克拉底到亚里士多德和亚力山德里亚的一般解剖学家；历史学从希罗多德到修昔底德与波利比阿；逻辑学、政治学、道德学、美学，从柏拉图、色诺芬、亚里士多德到斯多葛学派和新柏拉图学派。如此醉心于思想的人不会不爱好最崇高的思想，概括宇宙的思想。………在别的地方，机能的天然的平衡受到文明破坏；文明总是夸大一部分机能，抑制另一部分机能；把现世为来世牺牲，把人为神牺牲，把个人为国家牺牲。文明造成印度的托钵僧、埃及与中国的官僚、罗马的法学家与税务官、中世纪的修士、近代的人民、被统治者、资产阶级。在文明的压力之下，人有时胸襟狭窄，有时兴奋若狂，或者两者兼而有之。他成了一架大机器中的一个齿轮，或者觉得自己在无穷的宇宙中等于零。在希腊，人让制度隶属于人，而不是人隶属于制度。他把制度作为手段，不以制度为目的。他利用制度求自身的和谐与全面的发展；他能同时成为诗人、艺术家、批评家、行政官、祭司、法官、公民、运动家；他锻炼四肢，培养趣味，集一二十种才能于一身，而不使一种才能妨碍另外一种——这解释了什么叫做身心的平衡。他可以成为士兵而不变做机器，成为舞蹈家、歌唱家而不成为舞台上的跑龙套，成为思想家和文人而不变做图书馆和书斋中的学究……为神明举行赛会而不受教条束缚，不向一种超人的无穷的威力低头，不为了一个渺茫而无所不在的神灵沉思默想。仿佛他

们对于人与人生刻画了一个感觉得到的分明的轮廓……这才是真实的人，一个有思想、有意志、又活泼敏感的身体；这才是真正的人生，在呱呱而啼的童年与静寂的坟墓之间的六七十年寿命。我们要使这个身体尽可能地敏捷、强壮、健全、美丽，要在一切行动中发展这个头脑、这个意志，要用精细的感官、敏捷的才智、豪迈活跃的心灵所能创造和欣赏的一切的美，点缀这个人生。”[①]

好了，丹纳这段长长的话足以证明科学、艺术以及快乐原本是一家的，只不过由于近代以来学问的分工，才导致人的发展被局限在狭窄的领域内，不同学科不同领域之间才造成“隔行如隔山”，所谓“鸡犬之声相闻，老死不相往来”的局面！所以我们需要“返璞归真”，重新追寻古希腊人的全面发展和快乐学问的精神。《策略博弈》以大量趣味盎然的故事，将博弈论的精彩应用向我们娓娓道来，在博弈中玩游戏，在游戏里讲博弈，十分好玩。上个世纪初的大数学家希尔伯特教授曾说过，真正的数学大师是能够在乡间小道上向偶然遇见的农夫讲清楚什么是微分几何的。这一点也适用于博弈论和《策略博弈》。现在我们也可以说，迪克西特教授是博弈论圈中真正的高手，是能够通过讲故事把博弈论的基本道理向基本没有多少高等数学知识背景的读者们说明白的优秀学者。

迪克西特教授本人十分关心《策略博弈》中文版的翻译出版。他亲自为中文译本撰写了序言，其中还大量引用了包括《孙子兵法》和《三十六计》在内的中国经典。他还亲自回答了我们在翻译中所遇到的一些问题，并且通过电子邮件发来了《策略博弈》最新的勘误表。

本书的翻译出版是集体合作的成果。博士生董志强（现为华南师范大学副教授）翻译第 1 章、第 2 章、第 9 章、第 12 章；硕士生陈斌（现为香港大学经济金融学院博士生）翻译第 4 章、第 5 章、第 11 章；博士生田盈（现为重庆师范大学数学学院副院长、教授）翻译第 6 章、第 7 章、第 8 章；蒲勇健与博士生杨红合作翻译第 3 章，博士生杨红翻译第 18 章，蒲勇健翻译第 13 章、第 14 章；蒲勇健与博士生刘渝琳（现为重庆大学贸易与行政学院副院长、教授）合作翻译第 15 章；博士生雷勇（现为长江师范学院组织部长、教授）翻译第 16 章、第 17 章；硕士生刘俊希与博士生刘渝琳合作翻译第 10 章，并且负责全书的初步校对。王宝来、冯丽君、顾晓波、胡安荣通读了全书，并提出了宝贵意见，最后由蒲勇健对全书进行终校。

---

① 丹纳著，傅雷译：《艺术哲学》，下册，350～358、361～362 页，天津，天津社会科学院出版社，2007。

中文版序

为什么要向中国人讲授博弈论呢？全世界除中国以外的其他地方显然应该从中国的历史和文献中汲取策略方面的智慧。的确，在美国的商业经济活动中，许多富有进取心的人士早已把《孙子兵法》作为了他们的行动指南。孙子所著的另外一部作品——《三十六计》[①]，为策略的使用提供了更加富有见地的阐述。该书对于诈术有着如下的指引：

1.“暗度陈仓”——偷偷地沿着陈仓通道突破楚河汉界（兵分两路向敌人发起进攻。第一路是直接的进攻，它是公开的并且会受到敌人防备的一种进攻。第二路是间接进攻，是居心叵测的，使得敌人未曾预期到且在最后一刻导致其分兵抵抗，从而将其引向混乱和灾难）。

2.“假痴不癫”——为了吃掉一只老虎而佯装成一头小猪（假装是一个傻瓜、醉鬼，或者是一个疯子，制造出有关你意图和动机的混乱状态。将你的对手加以诱骗至其开始低估你的能力，直到其过度自信且失去戒心。此时你就可以发起攻击了）。

① 原文如此，作者有误。《三十六计》不是由孙子所著。《三十六计》是根据我国古代卓越的军事思想和丰富的斗争经验总结而成的兵书。“三十六计”一语，先于著书之年，语源可考自南朝宋将檀道济（？—公元436年），据《南齐书·王敬则传》：“檀公三十六策，走为上计，汝父子唯应走耳。”意为败局已定，无可挽回，唯有退却，方是上策。此语后人赓相沿用，宋代惠洪《冷斋夜话》：“三十六计，走为上计。”及明末清初，引用此语的人更多。于是有心人采集群书，编撰成《三十六计》。但此书何时由何人所撰已难确考。——译者注

书中还有找出有关敌人意图和计划的信息的策略。

3. "打草惊蛇"——通过拍打其周边的草丛来惊动蛇（当你不能获知对手的策略时，可发起一场直接的但是迅速的攻击，然后观察你对手的反应。敌人的行为会显示出他们的策略）。

还存在一条有关调整价值的计谋：

4. "连环计"——绝不要单纯地依赖某一种策略（重要的是，应该同时能够使用多个可采用的策略。在一个总的计划中保持不同的方案处于可运行的状态，在这种情况下，倘若任何一个策略失败的话，你就仍然有其他一些储备好的策略可以使用）。

最后，把眼光放得长远一些也是重要的。

5. "李代桃僵"——用李子树取代桃树（丢掉芝麻捡西瓜。有些时候你必须放弃短期目标而追求长期的目标）。

最后一句格言与策略思维的另外一个重要方面相联系。策略性互动很少像战争那样是零和博弈的，其中一定要有赢家和输家。更为普遍的是，博弈往往是非零和的，并且双方都能够在长期的持续性合作中有所收获。所以，我们可以在《论语》中找出先贤所说的一条黄金法则："己所不欲，勿施于人。"孔夫子也承认下述策略的困难：

子贡曰："我不欲人之加诸我也，吾亦欲无加诸人。"

子曰："赐也，非尔所及也。"

所以，世界其他地方有许多东西都需要从中国丰富的策略思维宝库中学习。但是，我们认为，一部博弈论的中文译本也有其存在的理由。理论所要做的是将各种各样零散的例子整合成一种有着内在一致性的框架和形成一种分析的方法。当人们掌握了某种理论时，就可以把成千上万的例子加以归类，并且还会具有举一反三的潜力。现代博弈论就完全能够做到这一点，哪怕是包含在格言和谚语中的大量的有关策略的丰富智慧，都能够通过博弈论语言加以表述和分析。我们希望这本书将会有助于中国的读者整理他们先前的阅读知识并且以这种方式再度体会体会，以进一步改进他们的运用策略的技巧。

根据在美国和其他地方所遇到的许多才华横溢的中国学生的杰出表现，我们相信中国会在博弈论领域里作出新的重要贡献。我们也希望这本书能够作为一本入门教程为未来的学生和研究人员提供一个基本的起点。我们非常感谢重庆大学的蒲勇健教授和他的学生们将本书翻译成中文，还要感谢中国人民大学出版社出版该书的中文译本。

阿维纳什·迪克西特

苏珊·斯克丝

第二版前言

我们撰写这本书的目的是为了使得这样一件事情成为可能，即不需要经济学或政治科学或生物学知识的先期准备，而只需要最低程度的中学数学知识，就可以为一年级或者是二年级的大学生讲授导论或“原理”水平上的博弈论课程。我们对于所产生的反响感到十分的欣慰。在几年前，这种类型的任何课程都是几乎不可能存在的；现在许多教师都根据我们这本书开设了新课程，并且发现这种创新受到了学生们的广泛欢迎。本书还在中学生们的某些高级暑期项目中被使用。在最近几年里，博弈论不断被用于各个新的或旧的领域，并且还由于西尔维娅·娜莎（Sylvia Nasar）撰写的约翰·纳什（John Nash）传记《美丽心灵》（*A Beautiful Mind*），以及据此改编的由电影明星罗素·克鲁（Russell Crowe）主演的同名奥斯卡获奖影片而变得闻名遐迩。所以，人们对该学科的兴趣，以及对导论性课程的欢迎程度，也一定会与日俱增的。

我们从许多使用了第一版的教师和学生那里获得了极具价值的反馈信息。这样一些反馈信息，以及我们自己从本书中取得的教学经验，给我们提供了许多有关材料和阐释方面的改进思路。在此期间，博弈论理论本身也取得了某些进展，并且也出现了许多新的和十分有趣的应用。这里的第二版汲取了所有这些建议和最新发展的精华。

最为实质性的变化出现在由第3章到第8章构成（在第一版里是第3章到第7章）的第二部分里。我们重新撰写了这些章节，几乎是完整地详细描述了序贯和同时行动博弈的基本理论。我们重新安排了素材的顺序，以增进论述的流畅

性，删除或更换了某些被证明难度过高或直接相关性稍弱的章节，并且还对我们借以提出理论性概念的大多数例子进行了更改或改进。我们希望教师们将会发现调整顺序后的素材更便于使用，易于挑选。下面就是对上述章节的主要修改：

（1）对纯策略纳什均衡的处理现在被分为两章：第 4 章，它仅介绍只有有限个可选择策略的博弈；以及第 5 章，它介绍的是以连续性变量表达的策略（还包括其他的内容，诸如对纳什均衡概念的批评的一般性讨论和来自实验室的实验和现实生活博弈中的观察的证据）。这两章现在都把对纳什均衡的讨论作为一种信念与选择的体系加以涵盖。对于那些教授简短课程的教师们来说，可能需要省略掉第 5 章的部分内容。

（2）第 5 章用比第一版中第 7 章更长的篇幅讨论了纳什均衡的一个引人注目的替代概念，即可理性化（rationalizability）。这种讨论和在该章中前面部分的定价博弈为博弈论在产业经济学中的应用构建了一种较为充分的引导，许多教师都乐意在他们的课程中将其包括进去。

（3）对混合策略中的纳什均衡的处理也照样被分为了两章：第 7 章，处理的是零和博弈；以及第 8 章，它包含了非零和博弈和其他进一步的主题。这种分离强调了混合策略在这两种类型博弈中所扮演的不同角色。它也使得那些不喜欢非零和博弈混合策略均衡的教师们能够完全忽略掉那些材料。

（4）具有多重纳什均衡的协调博弈——纯协调、安全（assurance）博弈，以及性别战 的整个大类的扩展现在就通过“男孩子与女孩子约会”这样一个统一的例子加以完成了，该例子出现在许多章节里。我们希望该例子能够引起大学生们的兴趣。

（5）某些教师要求我们把处理序贯与同时行动博弈的顺序颠倒过来。我们没有采纳这种建议。我们的经验是，对于那些缺乏有关博弈论先期训练的学生们来说，策略的序贯行动概念比起同时行动概念要来得更加自然一些：想想“如果我这样干，其他人会做出什么样的反应呢，接下来又想想我下一个反应是什么”是比“我必须想想其他人是如何想我想的是什么”要更加自然一些的。所以，我们愿意基于这样一种倾向出发，表明它如何能够提炼为一种分析的方法（逆推法），然后继续走向同时行动所需要的新的思维及表达的方式。然而，那些喜欢别出心裁做事情的教师们将发现新的第 4 章和第 5 章是可以放到第 3 章之前讲授的：在第 4 章和第 5 章中只涉及少许对于序贯行动、博弈树以及逆推法的不重要的说明。

第三部分和第四部分中的章节没有进行大幅的修改，但是我们在所有的地方都提供了新的或者是改进过的例子，以及来自观察和实验的所有可能获得的新证据。在所有的章节里，我们都给出了几道新的练习题，并且把老的题目进行了修改并重新安排了顺序，这样可以给出一个更加完整和有更好难度顺序的序列。此外，除了在每一章的末尾列出关键术语外，我们现在还在书的最后给出了完整的词汇表①，并且对这些词汇加以了更规范严谨的定义。上述两部分最为值得注意的变化概括如下：

① 我们在翻译出版过程中将词汇表省略了。——译者注

（1）第三部分的顺序被重新编排了，使得关于不确定性与信息的章节首先出现；这就使得那些要较早地使用此类材料的教师变得较为方便一些。该章包括了关于“廉价磋商”（cheap talk）的新的一节，随后就是策略性行动，然后是关于重复博弈（囚徒困境）、协调博弈以及进化博弈的几章。最后的三章中使用了相同的案例和分析方法，这种新的编排顺序可使章节间的连接更加清晰，也可使这三个主题之下的课程材料组织变得更加流畅。但是，正如在第一版中那样，因为只存在相对少的交叉性参考材料，本书这一部分里的章节是可以以几乎任何顺序进行讲授的。

（2）在第四部分里，我们对有关投票和拍卖的章节进行了最为重大的修改。在“策略与投票”这一章中存在着一种对阿罗（Arrow）不可能性定理的应用的比较分析，以及对不完美信息情况下策略性投票的一种新的讨论。在“投标策略与拍卖设计”这一章里，我们把介绍美国联邦通讯委员会（FCC）的电磁波频段拍卖的一节换成了方兴未艾的网上拍卖。

自从第一版印刷出版以来，我们专门为教师们建立了一个网站，它带有一个教员指南和一个供讨论的论坛，主页是 www. wwnorton. com/college/econ/strategy/。我们鼓励教员们使用这个资源和在供讨论的论坛上交流他们自己的点滴心得和评论。

最后，特别要感谢那些在第二版的出版过程中给予我们帮助的朋友们。文森特·克劳福德（Vincent Crawford，加利福尼亚大学圣地亚哥分校），戴维·H·赖利（David H. Reiley，Jr.，亚利桑那大学），托马斯·普鲁沙（Thomas Prusa，罗格斯大学），以及格雷格·特兰德尔（Greg Trandel，佐治亚大学）都对此书新版的写作进行了评论并提出了改进建议。在诺顿公司里，埃德·帕森（Ed Parsons，我们原来的编辑）和杰克·里普切克（Jack Repcheck，我们目前的编辑）两人在改进和提升本教程过程中所做出的不懈努力也值得我们感谢。同样，我们要感谢帕蒂·齐默尔曼（Patty Zimmerman）对该书新版进行了审校，还要感谢戈登·布鲁姆（Gordon Brumm）准备了索引，也谢谢朱莉娅·海因斯（Julia Hines）为本书最终的出版所做的努力。还要感谢伊·陈·翁奇（Ee Cheng Ong，威尔士里学院）协助我们对最终手稿进行了复查。苏珊·斯克丝可能还要额外感谢她的丈夫，保罗·凡·马尔希里格特（Paul van Mulbregt），因为她在整个修改过程中得到了他的鼓励与支持，同时要额外感谢她的孩子们，爱弥莉（Emily）和皮尔斯（Pierce），因为他们经常成为有关威胁与许诺的策略性运用的新例子的来源。

阿维纳什·迪克西特

苏珊·斯克丝

2003 年 9 月

目　录

第一部分　导论与一般性原理

第 1 章　基本思想和范例 …… 3

1.1　何为策略博弈？ …… 4

1.2　策略博弈的一些例子和故事 …… 5

1.2.1　回击到何处？ …… 5

1.2.2　GPA 暗斗 …… 6

1.2.3　“我们未能参加考试，是因为我们轮胎爆了” …… 7

1.2.4　教授为什么这样狠心？ …… 9

1.2.5　危机边缘的室友与家人 …… 9

1.2.6　约会博弈 …… 11

1.3　学习《策略博弈》的策略 …… 12

第 2 章　如何思考策略博弈？ …… 14

2.1　决策与博弈 …… 14

2.2　博弈分类 …… 16

2.2.1　博弈中各方是序贯行动的还是同时行动的？ …… 16

2.2.2　参与人的利益完全冲突还是具有某些一致之处？ …… 17

2.2.3　博弈只进行一次还是重复进行？对手是否固定？ …… 18

2.2.4 参与人拥有完全或相同的信息吗？ …… 18
2.2.5 博弈规则是固定的还是可人为操纵的？ …… 19
2.2.6 合作协议是否具有强制力？ …… 20
2.3 一些术语和背景假设 …… 21
2.3.1 策略 …… 21
2.3.2 支付 …… 22
2.3.3 理性 …… 23
2.3.4 对规则的共同知识 …… 25
2.3.5 均衡 …… 26
2.3.6 动态与进化博弈 …… 27
2.3.7 观察与实验 …… 28
2.4 博弈论的用处 …… 28
2.5 后续章节的结构 …… 29
2.6 总结 …… 31

第二部分 概念与技巧

第3章 序贯行动博弈 …… 37
3.1 博弈树 …… 38
3.1.1 决策结、枝及路径 …… 38
3.1.2 不确定性与“自然的行动” …… 38
3.1.3 结果与支付 …… 39
3.1.4 策略 …… 40
3.1.5 博弈树的构造 …… 41
3.2 用博弈树求解博弈 …… 42
3.3 增添更多的参与人 …… 46
3.4 行动顺序的优势 …… 50
3.5 添加更多的行动 …… 50
3.5.1 井字棋 …… 51
3.5.2 国际象棋 …… 52
3.6 逆推实验 …… 56
3.7 幸存者游戏策略 …… 58
3.8 总结 …… 61
第4章 纯策略同时行动博弈（Ⅰ）：离散策略情形 …… 67
4.1 离散型策略同时行动博弈的描述 …… 68
4.2 纳什均衡 …… 69
4.2.1 纳什均衡概念的进一步解释 …… 70
4.2.2 纳什均衡：一个信念和选择的体系 …… 71
4.3 占优策略选择 …… 72

4.3.1 两个参与人都存在占优策略的情形 …… 74
4.3.2 只有一个参与人存在占优策略的情形 …… 75
4.3.3 重复剔除劣策略 …… 76
4.4 最优反应分析 …… 78
4.5 零和博弈的最小最大值方法 …… 79
4.6 三人参与的博弈 …… 81
4.7 纯策略博弈的多重均衡 …… 84
4.8 不存在纯策略纳什均衡的情形 …… 88
4.9 总结 …… 90
附录 几个一般性的定义 …… 97
第5章 纯策略同时行动博弈（Ⅱ）：连续型策略、（Ⅲ）：讨论和证据 …… 99
5.1 纯策略是连续型变量的情形 …… 100
5.1.1 价格竞争 …… 100
5.1.2 政党宣传拉票 …… 103
5.2 纳什均衡的经验证据 …… 105
5.2.1 实验室或课堂里的实验 …… 106
5.2.2 现实世界的证据 …… 108
5.3 对纳什均衡概念的批评性讨论 …… 110
5.3.1 在纳什均衡中对风险的处理 …… 111
5.3.2 纳什均衡的多重性 …… 113
5.3.3 纳什均衡对理性的要求 …… 114
5.4 可理性化 …… 115
5.4.1 可理性化概念的运用 …… 116
5.4.2 可理性化导出纳什均衡 …… 117
5.5 总结 …… 119
第6章 同时行动与序贯行动并存的博弈 …… 125
6.1 同时行动与序贯行动并存的博弈 …… 126
6.1.1 两阶段博弈与子博弈 …… 126
6.1.2 多阶段博弈的结构 …… 129
6.2 在博弈中改变行动顺序 …… 131
6.2.1 同时行动博弈变为序贯行动博弈 …… 131
6.2.2 行动顺序上的其他变化 …… 137
6.3 分析方法的改变 …… 138
6.3.1 用博弈树表示同时行动博弈 …… 138
6.3.2 用策略形式说明和分析序贯行动博弈 …… 139
6.4 三人博弈 …… 142
6.5 总结 …… 145

第 7 章　采用混合策略的同时行动博弈（Ⅰ）：零和博弈 …… 149
7.1　什么是混合策略？…… 149
7.2　不确定行为：混合行动使对手难以预料 …… 151
7.2.1　最优反应分析 …… 151
7.2.2　最小最大值法 …… 156
7.3　纳什均衡作为一个信念与反应系统 …… 159
7.4　当一个参与人有三个或更多纯策略时的混合 …… 160
7.4.1　一般情形 …… 161
7.4.2　例外情形 …… 163
7.4.3　一个混合策略占优的情形 …… 164
7.5　当参与人双方都有三个策略时的混合 …… 166
7.5.1　所有策略的完全混合 …… 167
7.5.2　存在一些未使用策略的混合均衡 …… 168
7.6　零和博弈中混合策略的证据 …… 170
7.7　实践中如何使用混合策略 …… 172
7.8　总结 …… 173
附录　概率与期望效用 …… 179
第 8 章　采用混合策略的同时行动博弈（Ⅱ）：非零和博弈、（Ⅲ）：一般性讨论 …… 188
8.1　不确定信念支撑的混合策略 …… 189
8.1.1　哈里会碰上萨莉吗？…… 189
8.1.2　懦夫 …… 193
8.2　有三个策略的非零和博弈 …… 195
8.3　混合策略均衡的一般性讨论 …… 196
8.3.1　弱均衡 …… 197
8.3.2　对手的无差异与防止被对手利用 …… 197
8.3.3　在零和博弈中混合策略的违反直觉的结果 …… 198
8.3.4　非零和博弈中另一个违反直觉的结果 …… 200
8.4　非零和博弈中混合策略的实证 …… 201
8.5　三个或更多策略中的混合：一般性理论 …… 202
8.6　总结 …… 207

第三部分　某些更为广拓的博弈与策略类型

第 9 章　不确定性与信息 …… 213
9.1　非对称信息：基本思想 …… 214
9.2　直接沟通或“廉价磋商” …… 215
9.3　信号传递和甄别 …… 219
9.4　激励报酬 …… 222

9.5 信号传递博弈中的均衡 …… 225
9.5.1 分离均衡 …… 227
9.5.2 混同均衡 …… 229
9.5.3 半分离均衡 …… 229
9.5.4 一般理论 …… 230
9.6 信号传递和甄别的一些例证 …… 233
9.7 总结 …… 235
附录 信息与风险 …… 241
第10章 策略性行动 …… 249
10.1 策略性行动的分类 …… 250
10.1.1 无条件的策略性行动 …… 251
10.1.2 有条件的策略性行动 …… 251
10.2 策略性行动的置信度 …… 252
10.3 承诺 …… 253
10.4 威胁与许诺 …… 257
10.4.1 威胁的例子：美日贸易关系 …… 257
10.4.2 许诺的例子：餐馆定价博弈 …… 260
10.4.3 威胁和许诺相结合的例子——中美政治行动 …… 261
10.5 一些额外的议题 …… 262
10.5.1 策略性行动何时起作用? …… 262
10.5.2 遏制对强迫 …… 263
10.6 建立置信度 …… 264
10.6.1 减少你行动的自由 …… 264
10.6.2 改变你的支付 …… 265
10.7 反策略性行动 …… 268
10.7.1 非理性 …… 268
10.7.2 切断通信 …… 268
10.7.3 开辟逃走的路 …… 268
10.7.4 消除对方想要维持声誉的动机 …… 269
10.7.5 腊肠策略 …… 269
10.8 总结 …… 269
第11章 囚徒困境和重复博弈 …… 275
11.1 基本博弈（复习） …… 276
11.2 解法Ⅰ：重复 …… 276
11.2.1 有限次重复 …… 277
11.2.2 无限次重复 …… 278
11.2.3 未知期限的博弈 …… 280
11.2.4 一般理论 …… 281

11.3 解法Ⅱ：惩罚与奖励 …… 283
11.4 解法Ⅲ：领导 …… 285
11.5 实验证据 …… 287
11.6 真实世界的囚徒困境 …… 288
11.6.1 政府招商竞争 …… 289
11.6.2 劳动仲裁 …… 290
11.6.3 生物进化 …… 291
11.6.4 许诺最低价 …… 291
11.7 总结 …… 293
附录 无穷级数求和 …… 300
第12章 集体行动博弈 …… 303
12.1 双人集体行动博弈 …… 304
12.2 大群体中的集体行动问题 …… 306
12.3 思想简史 …… 310
12.4 解决集体行动问题 …… 311
12.4.1 分析 …… 312
12.4.2 应用 …… 314
12.5 溢出或外部性 …… 317
12.5.1 一般情况的计算 …… 318
12.5.2 负溢出 …… 319
12.5.3 正溢出 …… 322
12.6 "救命!"：有混合策略的懦夫博弈 …… 325
12.7 总结 …… 327
第13章 进化博弈 …… 334
13.1 研究框架 …… 335
13.2 囚徒困境 …… 338
13.2.1 重复博弈的囚徒困境 …… 339
13.2.2 多次重复博弈 …… 343
13.2.3 进化博弈模型与理性参与人模型之间的比较 …… 344
13.3 懦夫博弈 …… 345
13.4 安全博弈 …… 347
13.5 物种间的相互作用 …… 349
13.6 鹰鸽博弈 …… 352
13.6.1 理性策略选择与均衡 …… 353
13.6.2 $V>C$ 时的进化稳定 …… 353
13.6.3 $V<C$ 时的进化稳定 …… 353
13.6.4 $V<C$：稳定多态型种群 …… 354
13.6.5 $V<C$：每个参与人都采用混合策略 …… 355

13.7 种群中含有三种表现型 …… 356
13.7.1 检验 ESS …… 357
13.7.2 动态分析 …… 357
13.8 部分一般性理论 …… 360
13.9 领地博弈 …… 361
13.10 合作及利他的进化 …… 362
13.11 总结 …… 364

第四部分 在特定策略情形下的应用

第 14 章 边缘政策：古巴导弹危机 …… 373
14.1 对事件的一个简要的叙述 …… 374
14.2 一种简单的博弈论解释 …… 380
14.3 考虑更多复杂性后的博弈 …… 381
14.4 概率威胁 …… 386
14.5 边缘政策（冒险主义）博弈的操作 …… 389
14.6 总结 …… 393
第 15 章 策略与投票 …… 398
15.1 投票规则和程序 …… 399
15.1.1 二元法 …… 399
15.1.2 相对多数方法 …… 400
15.1.3 混合方法 …… 401
15.2 投票悖论 …… 402
15.2.1 孔多塞悖论 …… 402
15.2.2 议程悖论 …… 403
15.2.3 颠倒悖论 …… 404
15.2.4 改变投票方式就改变了投票结果 …… 405
15.3 关于悖论的一般性思路 …… 406
15.3.1 布莱克条件 …… 407
15.3.2 稳健性 …… 408
15.3.3 强度排序 …… 409
15.4 投票的策略性操纵 …… 410
15.4.1 相对多数规则 …… 410
15.4.2 配对投票 …… 411
15.4.3 不完全信息下的策略性投票 …… 414
15.5 与操纵有关的一般性思路 …… 416
15.6 中位投票者定理 …… 417
15.6.1 离散的政治光谱 …… 418

15.6.2 连续的政治光谱 …… 420
15.7 总结 …… 422
第 16 章 投标策略与拍卖设计 …… 431
16.1 拍卖的类型 …… 432
16.2 赢家的诅咒 …… 434
16.3 投标策略 …… 435
16.4 维克里说真话规则 …… 437
16.5 全部投标者都要支付的拍卖 …… 438
16.6 如何在拍卖中进行销售 …… 440
16.6.1 风险中性的投标者与独立估计 …… 440
16.6.2 风险规避的投标者 …… 441
16.6.3 相关的估计 …… 442
16.7 其他需要思考的问题 …… 443
16.7.1 多个投标品 …… 443
16.7.2 系统缺陷 …… 444
16.7.3 信息泄露 …… 446
16.8 互联网上的拍卖 …… 446
16.9 拍卖理论与实践的其他读物 …… 448
16.10 总结 …… 449
第 17 章 谈判 …… 454
17.1 纳什的合作解 …… 455
17.1.1 数字实例 …… 455
17.1.2 一般理论 …… 456
17.2 可变威胁的谈判 …… 460
17.3 轮流出价模式一：总价值的损失 …… 462
17.4 实验证明 …… 464
17.5 轮流出价模式二：不耐烦 …… 466
17.6 在谈判中操纵信息 …… 470
17.7 多方参与和多项议题的谈判 …… 472
17.7.1 多项议题的谈判 …… 472
17.7.2 多方参与的谈判（多边谈判） …… 473
17.8 总结 …… 473
第 18 章 市场与竞争 …… 477
18.1 简单交易博弈 …… 478
18.2 核 …… 485
18.2.1 数字例子 …… 487
18.2.2 核的一些特征 …… 488
18.2.3 讨论 …… 489

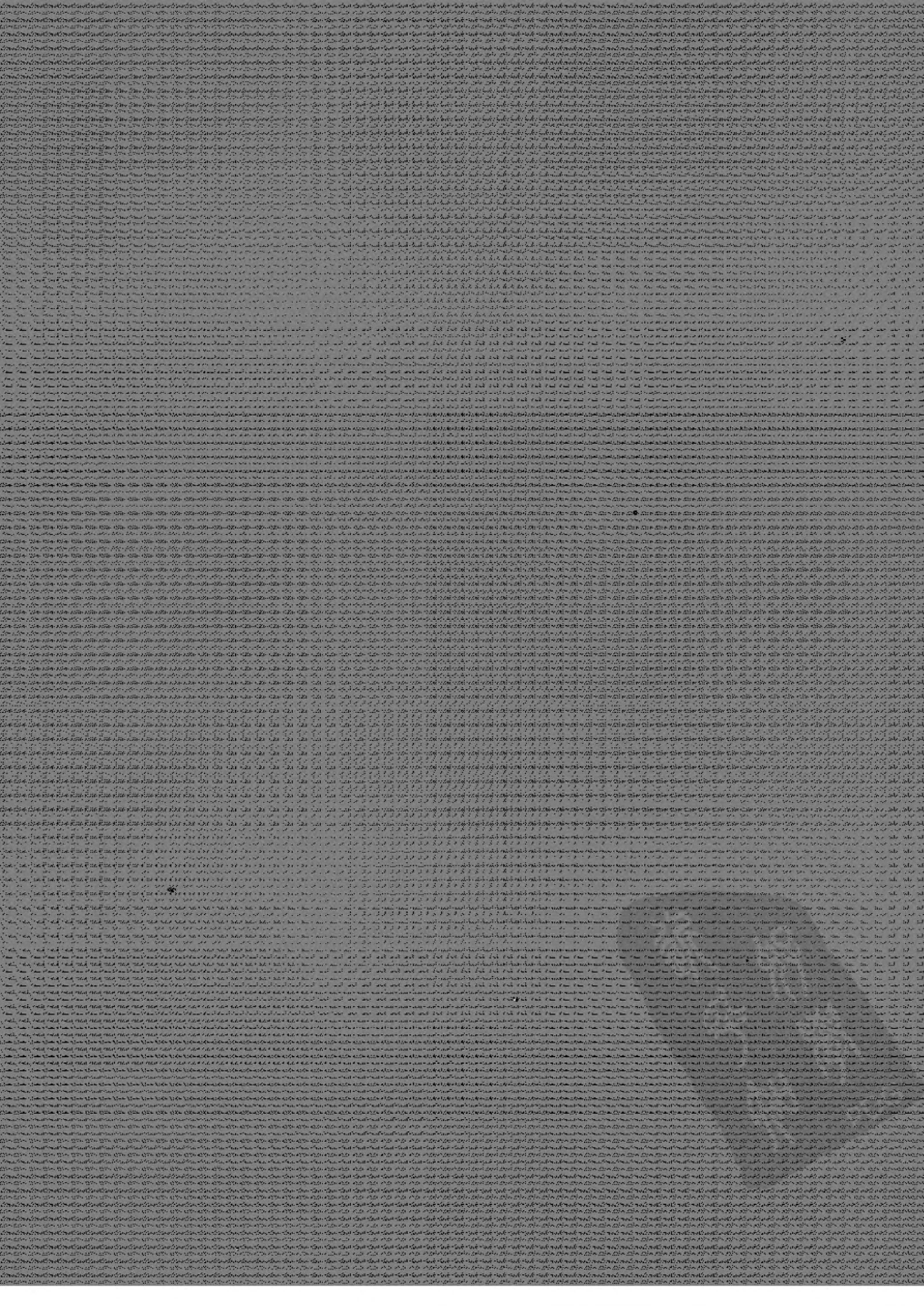
18.3 市场机制 …… 490
18.3.1 市场机制的特征 …… 492
18.3.2 实验证据 …… 493
18.4 夏普利值 …… 494
18.4.1 立法机构和委员会的权力 …… 496
18.4.2 联合成本的配置 …… 497
18.5 公平配置机制 …… 499
18.6 总结 …… 500

经济科学译库·策略博弈 经济科学译库·策略博弈 经济科学译库·策略博弈 经济科学译库·策略博弈

第一部分

导论与一般性原理





第 1 章　基本思想和范例

所有入门教材开篇都试图使读者相信该学科举足轻重，从而赢得他们的关注。物理科学和工程学号称自己是现代科技直至现代生活的基础；社会科学探讨重大的政府管理问题——譬如民主和税收；当你为物理科学、社会科学以及工程学而心力交瘁的时候，人文科学则声称它们可以使你恢复活力。那么，策略博弈，即通常所谓的博弈论，在这些学科中又有何重要地位？你又为何该学点博弈论呢？

我们将提供一个更切实际，且更符合个性关怀的学习动机。你随时都在博弈——跟你的父母、兄弟姐妹、朋友以及敌人，甚至跟指导你的教授。你可能已获得了大量的本能性技巧，而我们希望你会在本书中发现那些你已经学到的经验。我们将以这些经验为基础，进行系统的整理，以便你提高策略技巧并运用自如。在你的一生中，将有很多机会运用这些技巧，和你的老板、员工、爱人、孩子甚至陌生人进行这类博弈。

本学科覆盖了广泛的领域，类似的博弈在商业、政治、外交以及战争中经常进行——事实上，每当人们为达成协议或解决冲突而彼此互动时，博弈就发生了。洞悉这样的博弈，可丰富你对周遭世界的理解，并使你在纷繁世事中成为一个更优秀的参与者。

博弈论也有助于你学习其他科目。经济学和商学课程业已大量运用博弈理论思维。政治科学也正在迅速跟进。生物学已深受进化博弈概念的影响，而这些概念也已被引入到经济学中。心理学和哲学也与策略博弈的研究相互影响。可以

说，除了那些只探讨完全无生命物体的学科之外，博弈论业已成为一门为诸多学科提供思维方法和分析技巧的学问。

1.1 何为策略博弈?

游戏（game）一词可能传达了这样一种印象，即该学科无足轻重——它涉及的似乎只是赌博和竞赛之类的琐事，而世界上却充满了诸多更重要的事情，比如战争和商业，以及你的教育、职业生涯和人际关系。实际上，策略博弈（game）并不只是“一场游戏”（game）；所有这些重要议题都是博弈的例证，而博弈论有助于我们进一步了解它们。我们不妨以赌博和运动赛的例子作为开始。

绝大多数的游戏包含不同程度的机会、技巧以及策略。通过抛硬币来决定“要么倍赚，要么输光”是一个纯机会的游戏，除非你在作弊或抛掷硬币方面有超越常人的技巧。百码赛跑是一个纯技巧的游戏，尽管其中也可能混入一些机会因素，例如，某些参赛选手出赛时运气就是那么背。

策略是另一种类型的技巧。在赛场上，它既是打好比赛所需的心理技巧的一部分，又是对将体能技巧发挥到何种程度的一种算计。例如，在网球方面，你可通过练习发球（一发要用力而平稳，二发则需要旋转击球）和回击（用力、低、平且准确）来提高自己的动作技巧。策略上的技巧就是要知道球该往哪里发（远球或者是近网球）或者回击到哪里（底线球或斜线球）。橄榄球比赛时，你可通过跟防、围跑以及接传球来提高动作技巧。了解敌我动作技巧的教练通过场外喊话，指导球队最大限度地发挥本队的技巧，利用敌队的弱势。教练的算计组成了策略。橄榄球比赛的体能竞技是队员在场上进行的；而策略博弈则是在场外的指挥中心由教练或助理们进行的。

百码赛跑要求尽可能发挥体能技巧，因为它不会给你提供观察对手并做出反应的机会，因而无策略的用武之地。而长跑则的确需要策略，比如你是否应当领跑来确定跑步节奏，离终点多远时开始加速以摆脱其他对手，等等。

策略思维在本质上是你与他人的互动：他人也同时在进行着类似的思考。马拉松比赛中，你的对手基于对自身利益的考虑而试图阻止或诱导你领跑。网球赛中，你的对手试图猜透你的发球或回击落点；橄榄球赛中，敌队的教练将考虑你方的战术而采取最能打败你方的战术。当然，正如你必须考虑到其他选手的想法一样，别人也在考虑你所考虑的东西。博弈论就是进行这样的互动决策分析的科学。

若你三思而后行——你明白自己的目标和偏好，以及行动的约束和限制，并根据自己的标准以一种谋略方式选择行动以期得到最好结果——则可称你在理性地行动。博弈论为理性行为增加了一个维度，即与其他众多的理性决策者互动。换句话说，博弈论是关于互动情形下理性行为的科学。

我们不能说博弈论可以教给你精于对抗的秘诀，或确保你永不失败。首先，你的对手一样可以读到这本书，没有人可能永远获胜。其次，许多博弈都复杂而

微妙，而且绝大多数现实情形包含了大量的特异性反应和机会元素，因此不能指望博弈论为你的行动选择提供万无一失的方法。它所能做的，仅仅是提供一些思考策略互动问题的一般原理。在你谋划一个成功的策略之前，你必须把这些博弈思想和一些具体的计算方法补充到你所处的场合中。高明的策略家常将博弈理论和个人经验结合在一起，因此也可以说策略对局是科学，更是艺术。本书既会详述一般的科学思想，也会指出其局限性，并告诉你艺术性在何时会更重要一些。

或许，你已经从经验和直觉中知晓了对局的艺术，即便如此，你还是会发现学习这门科学将大有裨益。这门科学综合了大量的一般原理，这些原理在不同背景和应用中是共通的。若没有这些一般原理，则在每一个新的需要策略思维的情形下，你的思考都必须从头来过。在面对新的应用领域时，这样做是尤为困难的，比如，你从与父母和兄弟姐妹的博弈中学习到对局的艺术，但现在你却必须与生意场上的对手进行策略对抗。博弈论的一般原理为你提供了一个参照点。有了这样一个基础，你就可以更快更自信地在思考中抓住此种状况的具体特征，在行动中加入艺术成分。

1.2 策略博弈的一些例子和故事

如前所述，本书将从一些简单的例子开始。其中大多数例子来自你生活中可能遭遇的情形，在这些情形中策略非常重要。我们将针对每个例子指出它关键的策略原理，并在随后各章节中详细讨论；而在每一个例子结尾，我们将会告诉你可以从本书何处找到更详细的讨论。但请不要直接跳到详细讨论的章节。花点时间读完这些例子，可使你从整体上对策略和策略博弈有一个初步了解。

1.2.1 回击到何处?

最具水平的网球赛，是以下顶尖高手之间令人难以忘怀的对战：约翰·麦肯罗（John McEnroe）对伊万·伦德尔（Ivan Lendl），佩蒂·萨帕蕾丝（Pete Sampras）对安德烈·阿加西（Andre Agassi），以及马丁娜·纳芙拉蒂洛娃（Martina Navratilova）对克丽丝·埃弗特（Chris Evert）。回想 1983 年美国网球公开赛中埃弗特和纳芙拉蒂洛娃之间的决赛[1]，站在网前的马丁娜·纳芙拉蒂洛娃将球打向站在底线的克丽丝·埃弗特，而克丽丝·埃弗特正要回击，那她应当回击斜线球还是底线球？马丁娜·纳芙拉蒂洛娃应该预期短球而向前跑，还是预期穿越球而向另一边跑？

一般人的想法比较偏好短球。因为球越过网的距离较短，则留给对手的反应时间较少。但这不意味着埃弗特应该次次都放短球。因为如果她这样做，马丁娜·纳芙拉蒂洛娃就可以轻松断定落球点并准备回击，放短球就不容易取胜。为了提高短球的成功率，克丽丝·埃弗特会回击足够多的穿越球，让马丁娜·纳芙

拉蒂洛娃每次都得猜测球的落点。

同样，在橄榄球比赛中，当攻方只差一码就达阵时，直接冲刺就是所谓的致胜一击（percentage play），这是最常用的方法，但攻方在这种情况下偶尔还是必须传球，以免正中守方下怀。

因此在这种情况下，最重要的原则不是克丽丝·埃弗特应该做什么，而是她不应该做什么：她不能每次都用同一招，或者招数太有规律性。如果她这样做，马丁娜·纳芙拉蒂洛娃就会特意防守某个角落，而克丽丝·埃弗特的赢球机会就会下降。

杜绝有规律性地做一件事并不仅仅意味着避免在类似的情况下使用同一招数。克丽丝·埃弗特甚至不能在放长球和短球两种招数之间机械地切换，因为马丁娜·纳芙拉蒂洛娃会观察或识别这种模式（pattern），或发现其他的规律。克丽丝·埃弗特必须在每个特定的情况下随机地选择招数，以免被猜中。

包括电视上的体育评论员在内，大家都已熟知混合出招（mixing one's plays）的一般思想。但这一思想还有更丰富的内涵，需要更深入的分析。为什么放短球是惯用招数？球员放短球的次数应该占80%还是90%或者99%？在重要赛事中，策略是否应有所改变？比如，前述橄榄球达阵情形下，球员是否应在季度赛中传球而在超级杯赛中直接冲刺？在实际比赛中，球员应如何混合其出招？如果存在第三种选择（吊高球），情况又会如何？我们将在第7章和第8章讨论并解答这些问题。

电影《公主新娘》（*The Princess Bride*，1987）中，英雄韦斯特利（Westley）和恶棍维兹尼（Vizzini）之间的“斗智”（battle of wits）也反映了同样的思想。韦斯特利背着维兹尼往两杯酒中的其中一杯里下毒，而维兹尼决定谁饮哪杯酒。维兹尼对韦斯特利可能会把毒药放进哪一杯酒里进行了大量的循环推断，但是所有的推断本身是互相矛盾的。因为韦斯特利可以预期维兹尼的逻辑并选择把毒药投入另一杯酒里。反推过来，如果韦斯特利用特定的逻辑或规律选择某一杯酒，维兹尼就可以预期到并饮用另外一杯，让韦斯特利饮用剩下的有毒的那一杯。因而，韦斯特利的策略必须随机或非规律化。

这一幕也揭示了某些其他道理。在影片中，维兹尼输掉了游戏和性命。但影片后来显示，韦斯特利在两杯酒里都下了毒；在过去几年，他已经对这种毒药有了免疫力。所以维兹尼实际上是在一个致命的信息劣势下玩这场游戏。参与人有时可以应对这样的不对称信息；第9章将考察人们在何时可以以及如何应对信息的不对称。

1.2.2 GPA暗斗

假如你选了一门课程，课程成绩评定按某曲线分布进行。无论在绝对意义上你多么用功，都只有40%的学生可以得A，40%的学生得B。因此，你必须不仅需要在绝对意义上努力，而且相对于其他同学（其实在这里用“同班敌人”可能更恰当些），你还得更用功才行。全部同学都知道这一点，并且在第一次课后你

就组织大家开了个重要的会议，在会上全体同学皆同意大家都不要太用功。时间一周一周地过去了，稍微努力一点点以获得好成绩的想法逐渐占了上风。毕竟，其他人并不能观察到你的用功程度，对你也没有实际的约束力。而提高成绩等级的好处是实实在在的。所以你到图书馆去的次数更多了，熬夜的时间也更长了。

麻烦在于，每个人都在这样做。相较于大家都遵守约定的情况，现在你的相对成绩并不会得到改善。唯一不同的结果是，所有同学用功的时间都比约定的时间要多。

这是囚徒困境（prisoner dilemma）的例子。原本的故事是说，两个疑犯被隔离审讯并要求他们认罪。其中的 A 被告知："如果另一个疑犯 B 不认罪，你认罪，则可得到从轻发落；但如果 B 认罪，你最好是认罪，否则法庭将对你从重治罪。因此，不管 B 认不认罪，你都应该认罪。"同样，B 也听到同样的话，要求其认罪。面临这样的选择，A 和 B 都认罪了。但是，他们都不认罪对他们将是更有利的，因为警方并没有掌握真正有力的定罪证据。

你面临的情况与此类似。如果其他人都在混，你用功就可以取得更好的成绩；如果其他人也用功，你也只好用功，否则就会得到很糟糕的成绩。"囚徒"这个标签，贴到那些乖乖闭门用功的学生身上倒也是蛮合适的。

对于学校和教授，他们同样面临囚徒困境。每位教授都可以通过在成绩评定上高抬贵手来吸引学生修他的课，每所学校也可以通过同样的方式使学生更易于获得好工作或者吸引更好的申请者。当然，如果大家都这样做，其实谁对谁都没有相对优势；唯一的结果是，学生成绩水涨船高，分数差距缩小，让人难以根据分数判别学生的能力。

人们常常认为每个博弈中必有输有赢。囚徒困境却并非如此——所有的参与人皆可能是输家。人们天天都在进行（并输掉）这样的博弈，受损的程度从轻微不便到潜在灾难皆有可能。运动场上的观众站起来可获得更佳的视野，但若大家都站起来，其实每个人的视野并不比大家都坐着时更佳。超级霸权国家试图制造更多的武器来威慑竞争对手，但若双方都这样做，武力平衡没有改变；结果只是双方都耗费了本来有更好用途的经济资源，并且战争一触即发的风险也提升了。此类博弈的所有参与人都面临着巨大的潜在代价，因此弄清楚如何达成并维系互利合作非常重要。整个第 11 章都在研讨这样的博弈。

正如囚徒困境是潜在的双输（lose-lose）博弈，现实中也存在双赢（win-win）博弈。国际贸易就是一个例子。若每个国家都生产其具有相对优势的产品，每个国家都会从这种国际分工中获得好处。但为了实现贸易的全部潜在利益，必须对如何划分这块馅饼成功地进行讨价还价（bargain）。这一点同样适用于其他的需要议价的情况。我们将在第 17 章学习这些内容。

1.2.3 "我们未能参加考试，是因为我们轮胎爆了"

以下这则故事在本科生的电子邮件中流传甚广。当然也许只是一个编造的故事。笔者都曾经从我们的学生那儿听到这个故事。

两名交往甚密的学生在杜克大学修化学课。两人在小考、实验和期中考试中都表现甚优，成绩一直稳定在A。在期末考试前的周末，他们非常自信，于是决定去参加弗吉尼亚大学的一场聚会。他们玩得太尽兴，结果周日这天睡过了头，返校太晚，来不及准备周一上午的化学期末考试。他们没有参加考试，而是向教授编造了个悲伤的故事，说他们本已从弗吉尼亚大学往回赶，并安排好时间复习准备考试，但途中轮胎爆了。由于没有备用胎，他们只好整夜待在路边等待救援。现在他们实在太累了，请求教授可否允许他们第二天补考。教授想了想，同意了。

两人利用周一晚上好好准备了一番，胸有成竹地来参加周二上午的考试。教授安排他们分别在两间教室作答。第一道题目在考卷第一页，占了10分，非常简单。两人都写出了正确答案，心情舒畅地翻到第二页。第二页只有一个问题，占了90分。题目是："请问爆的是哪只轮胎？"

这个故事对今后参加聚会的同学有两大策略方面的重要教训。首先要意识到教授也可能是精明的博弈对象。他对学生的某些伎俩会心生疑窦并设法找出破绽。故事中的第二道考题就是针对学生的借口，教授用以找出他们的破绽的方法。两个学生应当事先想到这一点并备妥答案才行。参与人应当前瞻博弈中将来的行动，然后倒推出当前的最优行动，这是策略的一般原理。第3章将阐述这一原理。特别是在第10章，我们也会用到这个原理。

不过，学生要预见教授的此类破解妙招似乎不大可能。毕竟，教授看穿学生借口的经验，远比学生制造借口的经验要老到得多。如果两人没有事前准备，他们能够独立地编造出相互一致的谎言吗？如果每人都随机地选择一只轮胎，那么两人选同一只轮胎的几率只有25%（为什么?）。他们有没有更好的应对之策？

也许，你会认为副驾驶座一侧的轮胎最易爆，因为铁钉、碎玻璃更可能出现在路边而不是路中间，而这一侧的前轮会先遭遇它们。你也许认为这是一个不错的逻辑，但是这并不足以产生一个好答案。因为重要的不是候选项本身的逻辑，而是与你的朋友作出相同的选择。因此，你需要考虑你的朋友是否会使用同样的逻辑，跟你一样认为这个候选项比较好。但即便如此也无法结束推理过程，你的朋友会认为这个候选项对于你来说是明显较好的吗？等等。关键不在于某个候选项明显较好或合不合逻辑，而在于它对于对方来说是否是明显较好的候选项，双方又是否相信它对于对方来说是一个明显较好的候选项……换句话说，关键是特定情况下对参与人做何选择的预期之收敛（convergence of expectations）。参与人得以成功协调所依赖的这种共同预期策略称为聚点（focal point）。

所有此类博弈的结构本身并不能创造出这样的预期收敛。在有些博弈中，聚点可能因策略的机会因素或参与人共享的某些知识和经验而存在。例如，若因某种原因副驾驶座被称作"杜克位"，那么这两位杜克大学的学生很可能选择它而无须事前沟通。或者，若所有轿车的驾驶座都漆成橙色（为了安全，让迎面而来的车更容易看见），那么两位普林斯顿大学的学生很可能会选择该侧轮胎，因为橙色是普林斯顿大学的代表色。如果没有这些线索，双方根本不大可能有协调的默契。

我们将在第4章更详细地讨论聚点。在本节结束之际，我们只想指出当学生

在教室中被问及这个问题时，超过 50%的学生选择驾驶座那一侧。他们通常说不出所以然，只是说那好像是一个明显较好的答案。

1.2.4 教授为什么这样狠心?

许多教授都立有铁规，拒绝给补考机会，拒绝接受迟交的作业和学年论文。学生们认为教授这样做简直就是铁石心肠。而策略分析则给出了完全相反的解释。绝大多数教授都颇为仁慈，也愿意给学生合理的喘息机会，并接受学生所有的合理借口。问题在于判断何为“合理”。要区分类似的借口是很困难的，要一一确认真相也是不大可能的。教授很清楚自己可以给学生一次方便，但他也清楚这样做很危险。一旦学生发现教授“心太软”，他们就会拖得更久，编造更荒唐的借口。截止日期将失去意义，考试也会因缓考和补考而被弄得一团糟。

避免危险的办法通常只有一个，那就是绝不越雷池半步。拒绝接受任何借口是接受所有借口之外的唯一可行方法。教授可通过事先约定“借口免谈”的策略，避免做出让步。

但是，一个慈悲为怀的教授如何坚持这样一个铁石心肠的约定？他必须找到某些方法使其拒绝坚实可信。最简单的办法是拿管理办法或学校规定做挡箭牌。“我也想接受你的理由，但是学校不允许我接受。”这样既充当了好人，也使自己摆脱了无可选择只好破例的境况。当然，这些规则也可能是教授们集体订立的，但一旦制定则个别教授就不能在任何特殊情况下破例。

如果学校没有这方面的规定，教授也可以自己来做这样一个约定（commitment)。例如，一开课就明确而坚定地宣布政策。一旦有个别学生请求破例，教授就可以运用公平原则，即“如果我对你破例，那我就得对所有的学生破例”。或者，教授也可以通过几次毫不让步建立起严厉的声誉。对于他来说，这可能不是令人愉快的事，也可能违背他的本意，但这在其漫长的教学生涯中是有好处的。如果一个教授被认为非常严厉，就没有学生想用借口搪塞他，而他也就减少了拒绝学生的不愉快。

我们将在第 10 章仔细地学习约定及相关的策略，比如威胁（threats）和承诺（promises)。

1.2.5 危机边缘的室友与家人

假设你与几个同学同住在一套公寓房，发现洗洁精、卫生纸、麦片、啤酒和其他用品已经快用完了。大家同意平摊费用，但去商店购买得花费时间，你是自己花时间去呢，还是希望别人去以便自己有更多的时间学习或休息？你是上街买肥皂，还是待在房里看肥皂剧?[2]

在很多时候，这种等待博弈（waiting game）会僵持一段时间，直到有人受不了没有某种东西（通常是啤酒）而做出让步，花时间往商店跑一趟。这种等待

博弈有时也导致室友间相互争吵，甚至翻脸。

可以从两个角度来看待这种博弈。其一，每个室友皆有两种选择：采购或等待。对你而言，最好的结果是其他人采购而你待在家里面；最坏的结果是你去采购而其他人可以更好地利用时间。其二，如果大家都去采购（放学途中顺便购买，但彼此互不知晓），则可能产生不必要的重复购买，或者购买了太多的易腐品；如果没人去采购，则有可能带来严重不便甚至大麻烦，例如厕所的卫生纸刚好在紧急时刻用完了。

这很像美国年轻人过去常玩的懦夫博弈。两个年轻人各自驾车驶向对方，谁先扭转方向盘躲闪谁就是输家，而坚持直驶的人就是赢家。我们将在第 4 章、第 8 章和第 13 章进一步分析懦夫博弈。

同样情形的一个更有趣的动态视角被称为消耗战（war of attrition)。每个室友皆试图比别人等待得更久，希望其他人先失去耐心。这一过程中，风险在升级，因为一些重要物品可能会被用完，从而造成极大的不便甚至引起争执。每个人都任由风险被提升到各自忍耐的底线，忍耐力最差的人就输了。每个人都眼看着并任由情况发展到危险边缘，因此这种博弈中的这种策略被称为边缘政策(brinkmanship)。这是懦夫博弈的一个动态版本，它包含的可能性更丰富、更有趣。

笔者之一（迪克西特）有幸在某个周六的晚宴上曾目睹边缘政策博弈的一个精彩例子。开饭前，客人们正在客厅等候，这时主人家 15 岁的女儿出现在门边说："爸，再见。"父亲问："你要去哪里？"女儿回答："外面。"沉默了漫长的数秒钟后，主人家说："好吧，再见。"

在策略观察家看来，这一幕还可以不同的情形上演。主人也许会问："跟谁出去？"而女儿可能回答："朋友们。"父亲可以拒绝让女儿出去，除非她说清楚究竟跟谁要去哪里。父女一方或双方可能在一连串对话后做出让步，也可能引发争执。

这对父亲和女儿都是场危险的博弈。女儿可能会在客人面前遭到惩罚或责骂，而争吵也可能会败了客人和父亲整晚的兴致。由于不能确信对方是否会让步以及在何时让步，或者是否会出现不愉快场面，所以每个人都需要掂量这种僵持可以推进到哪种程度。若父亲极力逼问女儿，而女儿总是敷衍作答，爆发家庭战争的风险就增加了。

从这个方面看，父女之间的博弈正如工会和公司管理层进行的劳动合同谈判，或者又如两个超级大国之间蚕食对方全球势力范围的斗争。每一方都不完全明了对方的意图，每一方都一步步地窥探，而每一步窥探都提升了彼此间灾难的风险。故事中女儿试探着未知的自由之限度，而父亲也试探着了解自己权威的极限。

这是边缘政策的一个典型例子，一场提升彼此风险的博弈。这种博弈有两种结局。第一种结局是，某个参与人达到其承受风险的极限而做出让步（故事中父亲很快做出了让步。如果这位父亲一向家教甚严，则女儿根本不敢让这种场面发生）。第二种结局是，双方皆不让步，则双方担心的风险最终将爆发，争执（罢工或战争）产生。故事中主人家的不快以"愉快"告终，尽管父亲让步而女儿得逞，但总比来一场争吵要好得多。

我们将在第 10 章完整地分析边缘政策；在第 14 章，我们将考察它的一个特别重要的案例——1962 年的古巴导弹危机。

1.2.6 约会博弈

约会时，你总是想展示自己个性中最好的一面，掩盖糟糕的一面。当然，缺点不可能一辈子隐藏。随着关系的发展，你可以克服这些缺点，或者希望对方将你的优点和缺点一同接受。你也深知，若没有良好的第一印象，关系就不可能取得进展，因为你已没有第二次机会发展彼此的关系。

当然，你也想了解对方的一切，包括优点和缺点。但是你也深知，如果对方跟你一样是个约会博弈的高手，他/她也会同样地展示最好的一面而隐藏最糟糕的一面。你会仔细思考所面临的情形，并力图发现哪些迹象代表了真正的高素质，哪些只是为了给对方良好的第一印象而伪装出来的。最邋遢的家伙在重要的约会场合也可以摇身变为绅士，但要整晚每个细节都流露出谦恭和礼貌却并非易事。鲜花是相对廉价的，赠送更贵重的礼物也许是值得的。倒不是因为内在价值方面的原因，而是它代表了一个人乐意为你奉献多少的可靠证明。礼物值多少"钱"在不同背景下有显著差异。对于一个千万富翁来说，一颗钻石可能比牺牲宝贵的时间来陪你或替你做事便宜得多。

你也应当意识到，你的约会对象同样会对你的行为传达出的信息甄别一番。因此你得采取能真正表明你具有高素质的行为，而不是那些谁都学得来的行为。

探询、隐藏和发现对方内心深处的想法，不仅在初次约会时很重要，在整个关系发展的过程中都很重要。下面的故事说明了这一点。

一对青年男女，住在纽约，分别租有公寓。他们的关系已发展到同居的地步。女人向男人提议放弃他租来的公寓。这位男士是一名经济学家，他向女人说明了一条经济学原理：有较多的选择终归是比较好的。他们分手的概率很小，但是只要有分手的风险，则保留第二套廉租公寓就还是有用的。女人对这种说法非常反感，立刻结束了这段关系。

听了这个故事的经济学家说，这恰好证实了选择面越宽越好的原理。但策略思维提供了一种迥异但更有力的解释。女人无法确认男人对关系的忠诚度有多高，她的提议是发现真相的一种精明的策略机制。语言表达的爱总是很廉价的，因为每个人都可以说"我爱你"。如果男人用行动实践诺言，放弃了廉租房，这将是爱情忠贞的有力证明。而他拒绝这样做实际上是给出了负面证明，女人结束这段关系是明智的。

这些例子都试图唤醒你曾经历过的一些重要博弈的经验。在这些博弈中，关键的策略问题是对信息的操纵，传达关于你的正面信息的策略称为信号传递（signaling）；诱导对手传达关于他们私下的真正信息（无论正面或负面）的策略称为甄别机制（screening devices）。女人要求放弃一套公寓的提议就是一种甄别机制，它将男士置于这样一种境地：要么放弃公寓，要么暴露出他缺乏诚意。我们将在第 9 章学习信息博弈，包括信号传递和信息甄别。

1.3 学习《策略博弈》的策略

我们用几个日常生活常见的策略运用的例子，说明了策略思维和策略博弈的基本概念。今后我们还会讲授很多类似的故事，以期你面临真正的策略情形时可以发现它跟这些故事的相通之处，帮助你正确决策。这是许多商学院采用的个案分析法，它是帮助你发掘并记住潜在的博弈理论的有力工具。不过，每一个新的策略情形都组合了许多变量，要涵盖所有的变量，需要举无数的案例才行。

另一种方法是关注例子背后的一般原理，并构造策略行动的理论，即，正式的博弈理论。这样做的目的是当你面临一个现实的策略情形时，可以辨认出哪些原理适用其中。这是学术性学科（比如经济学和政治科学）中经常使用的方法。这种方法的缺点是，理论表达非常抽象和数学化，缺乏足够的个案和例子，从而会使很多初学者难于理解和记忆理论，也不易做到理论联系实际。

不过，了解一些基本理论确有必要。它让你更深刻地理解博弈，以及为什么会出现那样的博弈结果。相对于仅阅读某些案例和只知道怎样参与博弈，掌握一些基本理论对你将更有帮助。若知道为什么，你就可以看透那些新的未曾预期到的情形，而一个只知道“怎样”依葫芦画瓢的人却不能看透。世界跳棋冠军汤姆·威斯纳尔（Tom Wisnell）说得好：“知道怎样去做的人可以打成平手，而了解为什么的人却可以赢得比赛。”[3]但是也并非所有的博弈皆是如此。有一些博弈对任何人都是绝路一条，无论这个人知识何等渊博。但是这句话包含了一条重要真理——知道为什么给人们带来的优势超过只知道怎样做所带来的。比如说，知道为什么可帮助你预见某些绝路，从而一开始就避免身陷绝境。

因而，我们将采用一条中间路线，综合上述两种方法的长处——案例（怎样做）和理论（为什么）。我们将按照基本原理来组织专题，每章讨论一条原理，这样你就无须自己从案例中推导原理。我们将使用形象的案例而不是抽象的方式来推导基本原理，所以每种思想的脉络都非常清晰。换句话说，我们的精力放在理论学习上面，但我们将通过案例而不是抽象推演来学习理论。

当然，这种方法也需要某些妥协。最关键的是，每个案例都旨在阐释博弈论的某些基本观念和原理，因而我们抛弃了诸多与原理不相关的细节。如果某些案例看起来不太真实，请多多包涵。可以确定的是，我们抛弃那些细节肯定是有理由的。

请允许我们再啰嗦几句。尽管这些旨在发展概念和理论框架的案例都经过精心挑选（甚至以牺牲某些现实性作为代价），不过，一旦理论建立起来，我们会特别重视将其与实际相联系。本书通篇都用实际或经验证据来检验理论解释现实的能力。或好或坏的答案将给你谨慎使用博弈论的信心，激励你构建更好的理论。本书也会在适当的时候详细讨论现实中制度是如何演进的，以解决博弈论所指出的问题。不妨留意一下第 11 章关于囚徒困境如何出现在现实中又如何得到解决的讨论，以及 12 章中对更为一般的集体行动问题所做的类似的讨论。最后，

第 14 章将考察古巴导弹危机中边缘政策的运用。这种基于理论的案例分析，充分考虑现实细节，并进行详细的理论分析，已经成为不同领域如商学、政治科学和经济史中的常用教学方法。笔者希望通过对军事和外交上一段小插曲的原创研究，可以向你生动地介绍这种方法。

我们的方法是从个别案例导出一般原理，而后与现实互相验证并运用于现实。为了一以贯之地应用这一方法，我们必须首先确定用以组织本书讨论的总的原则。我们将在第 2 章从不同策略问题（或概念）的若干重要方面来对博弈做出分类。每一个方面都将被区分为两种对立的纯类型（pure type）。例如，行动顺序（order of moves）方面就包括两种类型：每个参与人依次行动（序贯博弈）和同时采取行动（同时行动博弈）。现实中很少有博弈完全地符合这种分类，绝大多数都融合了两种极端类型的某些特征。但考虑到各场博弈中所涉及的概念及其所涵盖的若干方面，以及每个方面中不同纯类型的融合方式，每场博弈都可归于我们的分类之中。要决定在某一特定情形下如何行动，人们可以把他们从纯类型中学到的知识进行恰当的组合。

第 2 章构建起总的框架之后，随后各章将在此基础之上展开论述，针对博弈中每个参与人的策略选择和全部参与人的策略互动，发展出一些基本思想和原理。

【注释】

[1] 埃弗特在 1975 年美国网球公开赛上第一次获得冠军。纳芙拉蒂洛娃在 1983 年决赛中第一次获得冠军。

[2] 这个例子来自 Michael Grunwald 的“居家”（At Home）专栏，参见“A Game of Chicken,” in *Boston Globe Magazine*，April 28，1996。

[3] 引自 Victor Nieederhoffer，*The Education of Speculator*，New York，Wiley，1997，p. 169。笔者感谢宾夕法尼亚大学 Austin Jaffe 提醒我们注意这句格言。

第 2 章　如何思考策略博弈？

第 1 章给出了一些策略博弈和策略思维的例子。本章则试图用更具系统性和分析性的方法来阐明这门学科。我们挑选了博弈中的一些重要概念（或方面），每一个概念（或方面）中都包含了一对对立类型。就博弈中时间这一方面来说，分别有严格依次行动（序贯行动）和同时采取行动（同时行动）。我们将考虑到对处于对立状态下的纯类型进行思考时暴露出的问题，除刚才所说的时间方面的问题外，还有如博弈只进行一次还是重复进行，以及参与人是否了解其他参与人等问题。

在以后各章，我们将更详细地讨论各类博弈及博弈中的各方面，并说明在一些特定的应用情景中如何运用博弈分析。当然，绝大多数实际应用并非纯类型，而是不同类型的组合。

此外，每个应用均可能涉及两三个相关类型，因此必须把我们从纯类型中得到的知识用合适的方式进行组合。我们会通过应用背景说明如何进行组合。

本章也将介绍一些分析中要用到的基础概念和术语，比如策略（strategies）、支付（payoffs）与均衡（equilibrium），对博弈论的应用做简要的讨论，并对全书架构提供一个总览。

2.1　决策与博弈

当某人（或团队、企业、政府）决定如何对付他人（或团队、企业、政府）

的时候，他们的行为必然存在某些交叉效应（cross-effect）。一个人的行为势必会影响另一个人所获得的结果。当人们问乔治·皮克特（George Pickett，盖兹堡战役的指挥官）内战中南方军失败的原因时，他回答说："我想这与北方军有关系。"[1]

不过，行为互动要变成一场策略博弈，还需要一些因素，即参与人相互意识到这种交叉效应。如果你知道别人的行为会影响到你，你可以对他的行为做出反应，或先行一步预测对方将来的行为对自己产生的有利或不利的影响，甚至先下手为强改变其未来反应来增进自己的好处。如果你深知对方也了解你的行为将影响到他，那你也就知道他将采取类似的行动。反之亦然。这种对交叉效应的相互认识，以及双方据此认识而采取的行动，构成了策略中最有趣的部分。

策略博弈和决策是有区别的。**策略博弈**（strategic games），有时简称**博弈**（game）（因为我们不关心诸如纯机遇或纯技巧之类的博弈），是双方基于交叉效应意识的行为互动。**决策**（decisions）是无须考虑他人反应的场合下的选择行为。如果李将军（Robert E. Lee，他领导皮克特作战）料想北方军在他的炮火下已无还手之力，他选择发动攻击就是一个决策；如果他意识到北方军正严阵以待，则选择发动攻击就成为致命的博弈的一个部分。简而言之，一场博弈需要有两个以上的参与人（player），每个参与人都对他人的行动做出反应（或者每个参与人会考虑其他人会怎么反应），否则就不是博弈。

策略博弈在两个参与人正面交锋中最为突出，比如20世纪50年代至80年代的美苏军备竞赛，通用汽车和汽车业劳工联盟之间的工资谈判，坦帕湾海盗队和奥克兰侵略者队在超级杯赛上的火拼。相反，大量参与人之间的互动中似乎不会出现相互意识的问题。比如单个农夫的产出相对于全国或全世界的农产品产出来说只是沧海一粟，一个农夫决定多种或少种点庄稼对市场价格无甚影响，因此没有必要把农业生产当做策略博弈。这是经济学多年来的通行观点。少数巨无霸企业之间的对抗则适合看做博弈，如美国汽车市场曾一度被通用、福特和克莱斯勒所主宰。不过人们一般还是认为，市场供求是主导大多数经济互动行为的力量。

事实上，博弈论涵盖的范围更广。许多情形是以成千上万人参与的非个人市场开始的，最终演变成两个或几个人之间的互动。境况如此变迁的两大原因是：相互承诺（mutual commitment）和私有信息（private information）。

首先，考虑承诺。当你打算建造一栋房子时，你可以从本地数十家建筑商中挑选一家，而建筑商也可以从潜在的客户中进行选择。这显然是一个非个人市场。然而一旦双方达成协议，客户需要支付首付款，建筑商需要按房屋设计购买原材料。双方脱离了市场而被捆绑到一起，变成了双边关系。建筑商可以偷工减料或者拖拖拉拉，而客户可以推迟下一期付款。策略要素开始进入这种情形。他们当初订立的合同必须预见到个体的动机并详细写明与工程进度联系在一起的付款安排。此外，合同还需依现实情况随时做出修正，而这些调整又会带来新的策略要素。

其次，考虑私有信息。成千上万的农民为其机器、种子、化肥等寻求资金，也有许许多多的银行可以提供贷款。但这种借贷市场也不是非个人化的。有良好农作技能、愿意下更多工夫的农民，更有可能获得好收成并偿清贷款；低技能或

者比较懒惰的农民，可能收成不佳而无力偿还贷款。违约的风险显然因人而异。违约的并非所谓的“市场”这个整体，而是个体的借款人。因此每一家银行都必须将其与每一位借款人之间的关系看做一场单独的博弈，它会调查每位借款人的信用状况并要求有借款担保。借款的农民将设法让银行了解自己的信用，而银行也会寻求有效的方式来查证农民承诺的真实性。

同样，保险公司会下一些工夫来调查投保人的健康状况，或在火灾发生后核查是否有纵火迹象；雇主会评估员工的能力并监督其表现。更一般地，只要在交易中参与人拥有一定的影响交易结果的私有信息，则每一个双边交易都是一场策略博弈，即便由成千上万类似交易组成的大型市场也是如此。

总之，当每个参与人都影响着行为互动的时候，无论其影响力是来自参与人本身的重要性，还是由于承诺或私有信息的缘故将关系范畴缩小到了其中重要的参与人，我们皆有必要将这样的行为互动看做策略博弈。此类情形通常是商业、政治甚至社交场合中的常见现象而非特例。从而，研习策略博弈对于分析上述议题的学科领域来说非常重要。

2.2 博弈分类

博弈在许多不同的背景下发生，因而有许多不同的特质值得研究。我们可以简单地将这些特质归为几个类别或方面。在各个类别或方面中，我们可以确定两种纯的博弈类型，其他所有的博弈可视为此两种纯类型的混合。我们通过几个问题来对博弈进行分类，这样做将有利于你思考现实中面临的博弈。

2.2.1 博弈中各方是序贯行动的还是同时行动的?

下国际象棋的行动是序贯进行的：白棋先行，然后黑棋行，再白棋行，依此类推。相反，在石油开采权或广播频道竞标中，因为不知道竞争对手的标价，竞标者相当于同时写下各自的标价。大多数真实的博弈兼有这两种特征。厂商在研发新产品的竞赛中是同时行动的，但是每个厂商都对竞争对手的进展有或多或少的了解，并借此做出反应。美式橄榄球赛中，对抗双方的教练同时下达战术，但四分卫可以在观察到对方的阵势后改变战术，或要求暂停以便教练改变招数。

序贯行动（sequential moves）和**同时行动**（simultaneous moves）之间的区别很重要，因为这两种博弈需要不同的互动思维模式。在序贯行动博弈中，每个参与人都必须思考：如果我这样做，我的对手会怎样反应？你对未来结果的算计支配着你现在的行动。在同时行动博弈中，你得精明地猜测对手现在会如何行动；但你必须认识到，对手在进行其算计时也会猜测你现在会如何行动，同时他也知道你在进行着同样的思考……你们双方都必须在这个循环中找出自己的策略。

在随后的三章中，我们将探讨两种纯粹类型的情况。第 3 章考察序贯行动博

弈，这类博弈中你必须着眼于未来而当下采取行动；第 4 章和第 5 章的主题是同时行动博弈，这类博弈中你必须在“他认为我认为他认为……”的循环中寻找最佳策略。我们对每一种情况都会提供一些简单的思维工具——博弈树和支付表——并获得一些指导行动的简明规则。

研究序贯博弈也可以告诉我们，何时有先动优势，何时有后动优势。大致说来，这取决于博弈中承诺和灵活性的相对重要性。例如，市场对抗企业的经济竞争博弈中，如果一家企业决心要实施攻击性的竞争，因而吓阻了对手，这就存在先动优势。然而在政治争斗中，对某个问题抱着坚定立场的候选人可能会给对手留下明确的攻击点，这样的博弈具有后动优势。

了解如何平衡各种考虑因素，也有助于你寻找到方法去操控行动顺序以获得优势。这反过来又引出了威胁和承诺等策略行动，对这些内容我们将在第 10 章予以探讨。

2.2.2 参与人的利益完全冲突还是具有某些一致之处?

下棋或球赛等简单博弈中，有赢家也有输家。一方所得就是另一方所失。同样，在赌博游戏中，也是一方所得即另一方所失，得失的总和为零。这种情况的博弈可称为零和博弈（zero-sum games)。更一般地，零和博弈的思想就是参与人的利益是完全冲突的。当参与人分割任一一定数量的利益时，这种冲突就会产生，无论这一利益是用什么单位加以度量的——码、美元、亩，或几勺冰淇淋。由于可得到的利益并不总是精确地为零，我们常用常和博弈（constant-sum games）这个术语代替零和博弈。本书有时会交替地使用这两个术语。

绝大多数经济博弈和社会博弈并非零和的。贸易或其他经济活动可以令每个人都从中获利。合作生产可以组合参与人的不同技能并产生协同效应，生产出比各方独立作业时更高的产量。不过各方的利益也并非完全一致；各方可以合作生产更大的饼，但如何分配这块饼的问题却可能导致冲突。

甚至像战争、罢工这类的博弈也不是零和的。核战争是只有输家的最明显的例子，而这种看法也由来已久。公元前 280 年，伊庇鲁斯（Epirus）的国王皮拉斯（Pyrrhus）在赫拉克里击败罗马，军队付出的沉重代价使他说出这样一句话：“我们再次战胜，但我们也输掉了一切。”这也是俗语“皮洛士胜利”（Pyrrhic victory）的由来。20 世纪 80 年代，企业收购（take over）的狂潮达到极点，竞争对手之间竞相出价的争斗导致接管成本不断攀升，以致胜利者之所得常常类似于皮洛士胜利。

现实中大多数博弈都存在竞争与合作的压力，如何处理竞争与合作的关系则常常成为博弈论分析中最有趣的部分。参与人知道，如果合作失败则对大家都没有好处，因此在划分领地和利益时，也会努力去化解冲突。战争或罢工中的单方面威胁，正是试图恐吓另一方接受其要求。

即便一个博弈对所有参与人来说都是常和的，但若存在三个（或更多的）参与人，我们就必须考虑其中两人联合起来对付第三人的可能性。这一可能使得对

联盟（alliance 或 coalition）的研究成为必要。本书会在以后的章节中考察并详释这些思想。

2.2.3 博弈只进行一次还是重复进行？对手是否固定？

相对于多次互动，博弈只进行一次从某些方面看来比较简单，但从另一些方面看来却更加复杂。试想，在单次博弈（one-shot game）中，参与人不必担心对手在未来的博弈中进行报复，或担心别人听说此事后对他今后的行动有所提防。因此单次博弈中的行动更可能是不择手段或残酷无情的。例如，一家汽车修理厂更可能对过路车而不是常客狠狠敲上一笔。

单次较量中，参与人彼此知之甚少。例如，他们的能力和优势何在，他们是否精于算计最优策略，是否存在可以利用的弱点，诸如此类。因此，在单次博弈中，严守秘密和出其不意是“好策略”的重要因素。

在持续关系的博弈中，考虑的重点正好相反。你有机会建立起个人声誉（强硬、公正、诚实、令人信赖等，视情况而定）以及进一步了解你的对手。

参与人可以在博弈过程中通过对瓜分共同利益的合理安排（轮流当赢家）更好地促进双方利益，也可以在未来的博弈中对骗子进行惩罚（以眼还眼，以牙还牙）。第 11 章讨论囚徒困境时，将探讨这些可能性。

更一般地，短期表现为零和的博弈在长期中可能会存在共同的利益。例如，每支橄榄球队都希望赢得比赛，但是他们都意识到激烈的竞赛更能吸引观众的眼球，这在长期对所有球队都有利。这就是为什么他们会同意拟订这样一个方案：让战绩较差的球队优先选择较佳的球员以平衡各队战斗力的差异。在长跑比赛中，选手们常常发展出许多合作策略，比如几名选手可以通过轮流领跑承受风阻来互相帮助；在快到终点时，合作关系即告瓦解，选手们各自奋力冲刺。

对于人们在日常生活中的策略行动，存在一些有用的经验法则可供参考。在一场冲突与合作并存的博弈中，你常常想大获全胜让对手灰飞烟灭，但又担心这样不给对方留后路，会让自己表现得像 20 世纪 80 年代最坏的雅皮士。这种情形下，你可能忽略了博弈的重复性或连续性。攻击性的策略可以带来短期的优势，但长期的副作用会令你代价更沉重。因此你应深入思考找到合作的机会，并据此改变你的策略。面对人生的复杂博弈，你会发现仁慈，正直以及“己所不欲，勿施于人”的黄金法则并非仅是古老的处世箴言，也是很好的策略。

2.2.4 参与人拥有完全或相同的信息吗？

国际象棋对局中，双方均了解目前的战况以及过去的棋步，并且也均了解对方以赢棋为目标。不过此种情况并非常见，绝大多数博弈的参与人均面临一些信息限制。信息限制的来源有两类：其一，在博弈的每一个行动时点上，参与人可能无法获悉决策所需的全部信息。这包括相关的外部环境——比如天气——的不

确定性，以及对方先前或当前的行动。这类情况称为**不完美信息**（imperfect information）。我们将在第 4 章提出同时行动博弈的理论，并在第 7 章的附录和第 9 章讨论不确定条件下做出抉择的方法。

其二，当一个参与人比另一个参与人了解更多信息时，阴谋诡计就会产生。这类情况称为**不完全信息**（incomplete information），或者更恰当的称呼是**非对称信息**（asymmetric information）。此种情形之下，参与人会试图推测、隐藏，有时还会传递其私有信息，这些做法是博弈和策略的重要组成部分。在桥牌或扑克牌局中，每位玩家都只能部分了解别人手中的牌。各方的行动（例如桥牌对局中的叫牌与出牌，扑克对局中已出牌张数和下注高低）都可向对手传递某些信息。每个玩家都试图不按牌理出牌以误导对手（在桥牌对局中则是要设法与对家取得默契），然而这样做的时候，必须意识到对手往往也熟知这一点，并会通过策略性思维来解读这些行动。

读者朋友们可能会认为，如果自己拥有信息优势，就应该尽可能保守秘密。但其实不然。比如说，你是一家制药企业的 CEO，贵公司正与对手竞争开发新药品。当贵公司的研究人员获得了巨大进展时，你应当让竞争对手知道这个事实，希望他们知难而退。在战争中，交战各方均希望保守其战术与部署的秘密；但在外交中，爱好和平的国家总希望其他国家理解其善意。

一个普适的原则是，人们应当有选择性地披露自己的信息——披露好的信息（即令对方的反应有利于你的那一类信息），严守坏的信息（即可能对自己不利的信息）。

还有一个问题值得深思。你的对手在策略博弈中是理性的，并且知道你也是理性的。他很清楚你有时会虚张声势或者制造谎言。因此，他们不会平白无故地相信你对自己的进展或能力所做的宣告。除非证据确凿，或者你的行动可以有力地证明你宣告的信息，才能说服他们。这类由信息优势方（more-informed player）刻意采取的行动称为**信号**（signals），运用信号的策略称为**信号传递**（signaling）。反过来，信息劣势方（less-informed player）也可以制造出某些情形，让对手在该情形下如实地透露出其本身的某些信息。这种策略称为**甄别**（screening），其运用的方法称为**甄别机制**（screening devices，“screening”一词是作为筛选或分离的意思使用的，而不是隐藏信息的意思）。不妨回忆，1.2.6 小节的约会博弈中，女主人公就是通过甄别机制来检验出男人对感情的忠诚度的。她提出让两人放弃两套廉租房中的一套，如果男人足够忠诚，他会通过放弃自己房子的行动，来传递对爱情忠贞不贰的信息。

现在我们可以明白了，当参与人拥有不同信息时，如何操纵信息本身就是一场博弈，甚至比后续的博弈更重要。这种信息博弈随处可见，在日常生活中扮演着很重要的角色。本书将在第 9 章对此类博弈给予更详细的讨论。

2.2.5 博弈规则是固定的还是可人为操纵的?

国际象棋、扑克游戏或者体育竞赛，其规则是既定的，所有参与人均必须遵

守，无论这些规则看起来多么专断和古怪。但是，在商界、政坛以及日常生活的许多博弈中，参与人却或多或少可以自行拟订一套规则。比如，在家庭生活中，父母总是试图定出家规，而孩子们却总是设法改变或抗拒；在立法院，法案的进程规则（包括修正案和主要提议的表决顺序）是固定的，但是议程安排这场博弈——也就是决定哪一件修正案要先提出表决的博弈——却可以被人为操纵，而这也正是最能施展政治手腕和权力的场所。本书将在第15章讨论相关主题。

此类博弈本质上是订立规则的“赛前博弈”（pregame），你的策略手段必须在事前施展；而后续的博弈则比较机械化，你甚至可以交由他人全权代理。而如果你在浑然不知中错过了赛前博弈，你将会发现自己其实在比赛开始之前就已经输了。许多年来，众多的美国企业就是这样忽视外国企业的竞争力提升，而最终付出了代价。石油大亨洛克菲勒是一个例外，他就是只参与自己可以制定规则的博弈而获得成功的。[2]

在讨论策略行动（strategic moves）（譬如威胁与承诺）时，改变规则与遵守规则的差异显得尤其重要。

如何强化己方威胁和承诺的可信度，如何降低对方威胁的可信度，基本上与制定规则的赛前博弈密不可分。因为参与人必须在规则确定后的博弈中实现自己的威胁与承诺。本书第10章讨论的策略行动本质上就是一些用以操纵博弈规则的手段。

但是，若制定规则的赛前博弈是真正重要的博弈，那么赛前博弈的规则又是如何定出来的呢？通常，这与参与人本身的能力有关。在商场角逐中，企业可以先下手为强来改变整个博弈，例如通过扩建厂房或扩张广告的行动，让自己在今后与对手的价格较量中赢得先机，但究竟哪一家企业可以做得更好，这得视企业可用于投资或广告的管理或组织资源而定。

当然，参与人有时可能并不了解对手的实力。这往往让赛前博弈成为一种不平等信息博弈，参与人必须使用更精妙的策略，有时还会产生出人意料的结果。我们将在后面的一些章节中讨论这些问题。

2.2.6 合作协议是否具有强制力？

大多数策略互动都既有冲突也有共同利益。因此参与人应当走到一起，达成该如何行动的协议，以在最大化总体利益中平衡共同利益及在分配所得中平衡相互冲突的利益。协议的谈判可能要经过几个回合才能达成一致：在暂时决议后，大家再看看是否有更好的提议；一直到没有任何一方可以提出更好的意见为止，最终协议才会达成。本书第18章的核（core）概念包含了这样的过程和结果。但是，即便过程结束了，在协议执行中又常常会有新的问题。比如，所有参与人都必须遵守协议的条件，但若其他人都遵守条件，不遵守条件者往往可以得到更好的结果。但当每个参与人怀疑其他人可能背信弃义时，坚持合作行为就未免太愚蠢了。

如果所有参与人都在协议生效后立即在大家的监督下采取行动，那么这种合

作性质的协议有可能成功，不过这种立即实施的协议实在罕见。通常，参与人达成协议后就各自分散并独自行动。如果他们的行动可以互相观察，或者可由第三方——比如法院——来强制执行协议，则协议仍有很大希望继续维持下去。

然而在很多情况下个体行为都不能被直接观察，也无法由外力加以强制约束。如果缺乏强制力，只有在遵守协议对所有人皆有利时协议才不会破裂。主权国家之间的博弈、具有私有信息的博弈、其他无法可依或者因太过琐碎或成本太高法律不值得介入的博弈皆属此类。

事实上，这种集体行动的协议不具强制力的博弈，占了策略互动中的绝大部分。

博弈论使用两个专门术语来区分协议具有强制力和不具有强制力的情况。若协议对参与人行为具有强制力，则称此类博弈为**合作博弈**（cooporative game）；若协议没有强制力，个体参与人可根据其利益采取行动，则称此类博弈为**非合作博弈**（noncooporative game）。这已是标准的学术术语，不幸的是，一些人望文生义，往往误以为前者必然产生合作结果而后者不会产生合作结果。事实上，个人行动与众人的利益也可以是兼容的，尤其在重复博弈中更是如此。两种博弈的主要区别在于，非合作博弈中的合作行为只有当其符合个体参与人的个人利益时才会发生。这种由非合作性行动产生的合作结果是博弈论最有趣的发现之一。本书将在第 11 章、第 12 章和第 13 章讨论这一思想。

本书坚持标准的用法，不过仍要强调合作性和非合作性两个术语指的是行为实施的方式——前者采取集体行动而后者采取个体行动——而与博弈的结果无关。

正如我们在前面已提及的，现实中绝大多数博弈并没有充足的实施联合行动协议的外部强制力，因此本书以非合作博弈为主要分析对象。第 17 章对谈判的讨论，第 18 章对市场竞争的简要介绍，算是本书中的少数例外。

2.3 一些术语和背景假设

分析策略博弈之前，需先确定组成一个完整博弈的主要元素：参与人的可行策略、拥有的信息及其目的。按照前面所讨论过的博弈的若干方面对不同博弈的前两个元素进行考察，发现它们竟如此不同。但我们必须根据前面叙述的方式对不同博弈进行组合或构建。而后一个元素即参与人的目的最能引发新的思考。下面，我们将逐一讨论这些问题。

2.3.1 策　　略

策略（strategies），简单地说就是参与人可采取的选择。但即使是这样一个基础概念也需要深入讨论和阐释。如果一场博弈只能同时采取一次行动，则每个

参与人的策略正是这单次采取的行动。但是在序贯行动博弈中，后行动者可根据其他人（或自己）先前采取的行动而做出反应。因此每个参与人必须制定出一个完整的行动计划，比如“若对方采取 A 行动，则我采取 X 行动；若对方采取 B 行动，则我选择 Y 行动”。这种完整的行动计划就组成了博弈中的策略。

有一个简单的方法可以检验某项策略是否完整。一项完整的计划必须明确规定各种可能情况下将采取的行动。也就是说，若你把计划写下来留给其他的人去执行，然后自己去度假，这个人也可以照章行事，宛如你本人亲临博弈现场一样，不会出现意料之外的情况打扰你的休假。

在本书第 3 章中，我们就一些特殊情况深入探讨并运用此检验方式，使它更清晰地呈现在我们眼前。而现在你只需要记住：策略就是一个完整的行动计划。

这一概念类似于区别于战术的战略。战略一词（英文中“策略”和“战略”都是 strategy——译者注）常用于表示一个长期或大规模的行动计划，战术则指短期或较小规模的计划。例如，军方会为战争或一场大规模战役制定战略，而小型战役或战斗则由下层军官因地制宜采取战术。但博弈论根本不使用战术这一术语。策略这一术语涵盖了所有的情况，它既可意指一个完整的行动计划，也可意指一个特定的行动，如果这一单次行动就是我们所讨论的博弈全部的需要的话。

策略一词通常也可以用于描述一个人在相当长的时间跨度里所做的决策或一连串的选择，尽管我们还没有有目的地博弈的意识，也没意识到与他人的互动。因此，你也许已经有了某个生涯规划策略（career strategy）。当你开始赚到收入的时候，会做出储蓄、投资和退休策略规划。这种用法下的策略和我们观念中的策略具有同等的含义——它是对不断演变的环境做出反应的持续行动计划。唯一的差异是此处讨论的策略，其处境（即博弈）、环境在不断改变——由于其他有目的的参与人的行动而不停地改变。

2.3.2 支　付

当问及博弈中参与人的目的是什么时，绝大多数刚接触策略思维的新手认为是“赢得胜利”。但事实上答案并非如此简单。有些时候，获胜的“利润”（margins）很关键。比如在研发竞争中，倘若你的产品较竞争对手只有微小的优势，那么你的产品专利就面临着更大的竞争压力。也有些时候，参与人赢得的利益较小，输赢并不重要。更重要的是，只有极少数博弈是零和或非赢即输的，大多数博弈中参与人之间既有共同利益也有冲突。分析这类混合动机博弈要求更精明的算计而不是简单地以成败论英雄，我们还得比较合作与背叛的利益才行。

为比较博弈的所有可能结果，我们赋予所有参与人的每一个可行选择的组合一个数值。这个和每一个可能的结果联系在一起的数值就是参与人在该结果下的**支付**（payoff）。一般而言，对于参与人来说，更高的支付数值代表更好的结果。

在有些博弈中，支付只代表各种结果的好坏排序，最糟糕的结果标记为 1，

次糟糕的标记为 2，依此类推直到排列到最好的结果。而在另一些博弈中，支付数值是实际数字的表现，例如公司的收入、利润，电视台的收视率等。很多时候，支付数值只是合理的推测；然后我们进行敏感性检验，在合理的误差范围内改变这些推测，看我们的分析结果是否发生显著改变。

与支付有关的两个重要观念需要准确理解。其一，某一参与人的支付数值必须涵盖该博弈结果中他关心的一切事物。尤其需要指出的是，参与人不必是完全自利的，但利他倾向应已经包含在其支付数值中。其二，我们假设参与人面临随机结果时，其获得的支付是把各种可能状态下的支付值依其概率加权计算出的平均值。因此，假设参与人在 A 结果获得支付为 0，在 B 结果获得支付为 100，则当 A 发生的概率为 75%，B 发生的概率为 25%时，其支付为 0.75×0+0.25×100=25，我们称之为随机结果下的**期望支付**（expected payoff）。期望一词在概率论术语中有特别的含义，指的并不是预期可以得到的结果，而是数理、概率或统计意义上的期望。它指的是所有可能结果的平均值，其中每一个结果都以其发生概率作为求平均值时的权重。

上面第二点带来了一个潜在的难题。考虑一个赌钱的博弈，其支付就是金钱的数量。参看如下的例子：设想一个赌钱的博弈，其支付以金钱数量来衡量。若参与人有 75%的机会赚不到一分钱，有 25%的机会赚到 100 元，依前例其期望支付为 25 元。这相当于他在一场确定的可得到 25 元的博弈中得到的支付。换句话说，这种方式计算的期望支付 25 元与确定获得 25 元对于个人来说是无差异的。不过绝大多数人都是风险规避的，宁愿获得确定的 25 元而不是参与一场期望支付同样为 25 元的赌局。

我们简单地修正支付的计算方法来克服这个难题。我们放弃用金钱数量加总表示支付，而代之以用非线性函数（nonlinear rescaling）方式重新估算的数值。这就是所谓的预期效用方法，第 7 章附录中将对此加以详细说明。而现在，读者只需要知道可以把人们对风险的不同态度纳入分析。预期效用方法并非完美无缺，但却有极大用处，绝大多数博弈理论都以预期效用分析为基础。本书也采用这个方法，但在第 8 章第 3 节我们也会以一些例子指出这一方法中尚有待解决的难题。

2.3.3 理　　性

每个参与人均以获取最大支付为目标。但是参与人达成其目标的能力如何？这个问题的关键不在于其他对手追求其利益时会如何妨碍该参与人，毕竟这本来就是博弈的特性；关键在于，每个参与人是否能计算出对自己最有利的策略，并在实际的博弈中依计而行。

博弈论大多假设参与人精于算计并严格按照其最优策略行事，此所谓理性行为（rational behavior）假设。请注意理性一词的精确含义，它指的是每个参与人对所有可能结果（价值或支付）都有一套一致的排序标准来衡量哪个策略可以带来最大利益。因此，理性有两个重要的内涵：一个人对自己的利益完全了解，并

能完美地计算出何种行动可以最大化其利益。

认识到理性行为概念不包括什么也是同等重要的。理性行为并不意味着参与人是自私的；一个参与人可能对其他人的福利评价甚高，并将他人福利纳入自己的支付考虑之中。理性行为也并不意味着参与人是只局限于短期的思考者。事实上，考虑将来的后果也是策略思维的一个重要环节，某些短期看来不太理性的行为其实有其长远的策略意义。更重要的是，理性也不代表参与人必须与其他参与人持有相同的价值体系，敏感的人、古怪的人、道德高尚的人，他们的价值体系各不相同，理性只是表示参与人自己有一套一致的价值体系，并紧紧遵循这个价值链。因此，当参与人（在序贯行动博弈中）分析其他参与人的反应，或者（在同时行动博弈中）接连不断地推测如何推测时，他必须意识到其对手是根据其自身的价值体系决定其选择的。你不能将自己的价值体系或理性的衡量标准强加在他人身上，认为他们会在相同的情况下做出与你一样的行动。20 世纪 90 年代后期，许多评论海湾冲突的“专家”推测萨达姆（Saddam Hussein）会让步，“因为他是理性的”；这些专家忽略了萨达姆的价值体系与西方国家和评论家的价值体系是迥异的。

通常，每个参与人并不真正了解其他参与人的价值体系，这是现实中许多博弈存在不完全或不对称信息的部分原因。这类博弈中，试图了解对手的价值观，并试图隐藏或向对手传达自己的价值观就成了策略的重要部分。

博弈论假定所有参与人都是理性的。这个假设有多少合理性？博弈论运用这个假设的效果如何？在某种程度上，这个假设不完全真实。人们常常并不比别人更了解自己的价值体系，也不见得会事先逐一评估各种不同状况之优劣并铭记在心，除非他们已经面临必须做出选择的时候。因此，人们会发现，要一一细数自己与对手的各种选择可能带来的结果，并事先评价不同结果的优劣以决定最佳策略，实在是太困难了。即便他们都非常清楚自己的偏好，计算过程也会相当繁杂。实际生活中的大多数博弈非常复杂，而人们的算计能力也相当有限。比如在国际象棋之类的博弈中，众所周知，最优的策略只是有限步骤的计算，但是没有人能顺利完成这种计算，所以高明的棋招大部分还是一种艺术。

当参与人都是经常接触某一博弈的玩家时，理性假设才更贴近现实。他们得益于经历过或好或坏的结果，了解不同参与人不同的策略将如何影响结果，也知道自己技巧的高低。因此博弈的分析者可以预期他们在未经完整而精确计算下所做的选择会近似得到经过精确计算后的结果。我们可以将参与人视为一直做出最优决策的完美计算者（perfect calculator）。我们将在第 5 章提供一些实验证据，说明参与博弈的经验有助于参与人做出更理性的行为。

计算自己完整的最优策略时，把同样精于算计的对手的反应考虑进去，这样做的好处就是不会犯下让对手有机可乘的错误。在许多现实情形中，你可能在某些方面有特别的知识，而对手在这方面的理性比较缺乏，则你可以发现对方的错误并修订你的策略。我们会讨论一点这样的算计，但在博弈中，算计通常也是艺术的一部分，很难将其标准化为一套可供遵循的规则。当然，你必须留心对方是否假装他的技巧或策略很差，故意卖个破绽，欲擒故纵，先给你点小甜头，等你提高赌注时才大举反攻。这种情况下，参与人更安全的做法是把对手当做完全理

性的，替对手想好他的最佳策略。换句话说，人们的出招应着眼于对方的能力而不是缺陷。

2.3.4 对规则的共同知识

我们假设参与人在某种程度上对博弈的规则有共同的认知。在卡通片《花生》中有一幕，露西居然认为高尔夫球赛的规则容许身体接触，所以在查理挥杆时给了他一拐子。在博弈论中我们不允许这样的情况发生。

对“某种程度”的界定是很重要的。我们在前面已经了解到直接博弈的规则是可以被操纵的。但这仅仅承认了在更深层次上进行了另外一场博弈，在这场博弈中，参与人选择其表层博弈的规则。于是我们要问：这场更深层次的博弈是否有固定规则？例如在议会中，制定议程博弈遵循的是何种规则？答案也许是委员会主席们有权决定议程。那么委员会及委员会主席又是如何选出的？我们可以不断地追问下去。在某些基本层面上，这些规则由宪法、竞选技术或者一般的社会行为规范决定。我们要求所有参与人都了解基本博弈既定的规则，这是分析的重点。当然，这只是理想状况，现实中我们也许难以进行更深层次的分析。

严格说来，博弈的规则包括：（1）参与人名单；（2）每个参与人的可行策略；（3）每个参与人在所有参与人所有可能策略组合下获得的支付；以及（4）每个参与人都是理性的最大化利益追求者的假设。

博弈论不适用于以下情况：某参与人不知道是否有其他人参与博弈，不知道其他参与人可以选择的所有行动，不知道其他参与人的价值体系，或不知道其他参与人是否有意识地追求最大利益。然而，在现实的策略互动中，许多参与人以出其不意的方式，做出对方未曾预料的行动，结果获得更大的好处。这在战争史上有很多鲜明的例子，比如 1967 年以色列先声夺人地在地面摧毁埃及的空军力量，1973 年埃及出人意料地以坦克攻击苏伊士运河。

看起来，博弈论的严格定义忽略了策略行为的重要方面，但事实上没这么糟糕。我们可以修正理论，让每位参与人把对手可能的各种各样的戏剧化的策略作为发生概率较小的事件考虑进去。当然，每位参与人都清楚自己的可行策略。这样，博弈就变成了非对称信息博弈，第 9 章提到的方法可以对此进行分析。

共同知识这一概念本身也需要更深入的解释。某个事实或某种情况 X 要成为 A 和 B 两个人的共同知识，则他们各自知道 X 是远远不够的。每个人还都应知道对方知道 X；否则，A 也许会误认为 B 不知道 X，而在博弈中依此错误的认知采取行动。但是，A 也必须知道 B 知道 A 知道 X，反之亦然，否则 A 可能会错误地认为 B 不知道 A 知道 X，而错误地试图利用 B 的弱点加以攻击。当然，这个原则并非到此为止。A 应当知道 B 知道 A 知道 B 知道 X……如此反复，至于无穷。哲学家对这种无穷回归及其悖论有许多有趣的讨论，但对于我们来说，了解参与人对博弈规则有共同认识即已足够。

2.3.5 均　衡

理性参与人的策略互动最终会产生什么结果？我们将在**均衡**（Equilibrium）的框架下给出一般答案。均衡意味着每个参与人所采取的策略都是对其他参与人策略的最优反应。本书第 3 章到第 8 章将讨论博弈理论中的均衡概念，并在后续章节加以应用。

均衡并不是指事物不再变化。在序贯行动博弈中，参与人的策略是行动和反应的完整计划，随着博弈进行的每一步，局势都会不断变化。均衡也不表示所有的事情都臻至完美，所有参与人的理性策略互动有可能产生对所有人不利的结果，比如囚徒困境。不过我们会发现，均衡的思想通常是分析进程中相当有用的描述工具和概念。我们将联系具体的均衡概念来详细讨论这一思想。我们也将考察如何修正和加强均衡概念，以弥补它的缺陷，使它适用于完全理性计算下产生的非合作行为。

正如博弈经验的积累可以促成个体参与人的理性行为一样，参与人在博弈中经历的试错过程与获得的非均衡结果也能帮助其找到导致最终均衡的决策。我们将在第 5 章考察这个问题。

定义一个均衡并不难，而寻找一个特定博弈的均衡（即博弈的解）则要困难得多。本书将求解很多简单的博弈，这些博弈只牵涉两三个参与人，每人只有两三种可行策略，或轮流行动，或同时行动。很多人认为这正是博弈论研究的局限性，并因此认为该理论对现实中更复杂的博弈毫无用武之地，但事实并非如此。

由于人类的计算速度和耐心非常有限，只能解出只有少数参与人及少数策略的简单博弈。但电脑的计算速度和长度非常可观。许多超出人脑计算能力的博弈对电脑来说是很轻松的。电脑的计算能力已经可以处理很多商业或政治上的复杂博弈。即使是在国际象棋这种过于复杂而无法完整求解的博弈中，电脑已可达到人类最佳棋手的能力。在第 3 章我们将更多地讨论国际象棋。

求解复杂博弈的电脑程序正越来越多地涌现。Mathematica 及类似的程序包已经可以通过设定指令找出同时行动博弈中的混合策略均衡。加州理工学院教授理查德·D·麦凯尔维（Richard D. Mckelvey）与明尼苏达大学教授安德鲁·麦克伦南（Andrew McLennan）共同领导的国家科学基金项目 Gambit，正在创造一套综合指令，用以寻找序贯行动和同时行动博弈、纯策略博弈与混合策略博弈，以及具有不同程度的不确定性和不完全信息下的博弈的均衡。我们将在后面几章几个地方再次提到这个项目。该项目的最大优点是其程序皆被公之于世，大家可轻易地在 http：//hss. Caltech. edu/～gambit/Gambit. html 找到这些程序。

既然电脑可以帮助我们求解博弈，为什么本书还要详细地建立并求解一些简单的博弈呢？原因在于，了解这些概念是运用电脑求解的前提条件，练习一些简单的例题有助于加深对概念的理解。就像学习和使用算术的过程，首先得通过许多例题了解加减乘除的含义，掌握这些基本的概念后才能使用计算器或电脑进行复杂的运算。如果你不了解这些概念，就可能在使用计算器时出错，例如忽略了

先乘除后加减的规则，就会错把3＋4×5＝23计算成(3＋4)×5＝35。

因此，了解概念和工具这第一步非常重要。没有这一步，就无法知道如何正确地构造博弈让电脑求解，也不可能检查电脑的解答是否合理，更不可能在察觉结果不合理时回头修改条件，直到条件和计算能准确地抓住所研究的策略。因此，请认真研读本书提供的实例并解答后面的习题，尤其是第 3 章到第 8 章的习题。

2.3.6 动态与进化博弈

基于理性假设和均衡概念的博弈理论非常有用，但我们并不能完全依赖这样的博弈理论。当参与人是新手而或多或少缺乏进行计算并选择其最优策略的必要经验时，他们的策略选择和结果与以均衡概念为基础所预测的内容可能产生很大差异。

但是，努力做出好选择的观念也不能完全放弃，因为即便是计算能力很糟糕的参与人，也会为其利益而磨砺技巧，并从自己的经验和对他人的观察过程中逐步学习。我们应当允许这样一个动态过程，在此过程中先前被证明是较好的策略更可能在以后的出招中被采用。

博弈的**进化**（evolutionary）分析正是针对这一点的。进化分析派生于生物学中的进化思想。动物的基因对其行为模式有重大的影响，某些行为在当前环境中显得较为成功，在这种意义上拥有这些行为模式的动物更易繁衍并将其基因遗传给后代。进化稳定状态（evolutionary stable state）就是给定环境下经过数代进化过程的最终状况。

与生物进化类似，博弈中可假设极大化自身利益的理性参与人无法自行选择策略，而是由一个特定的策略“硬件”或“程序”来规定。其他参与对手也被设定采取相同或相异的策略，然后他们在博弈中各自得到支付。较为优势的策略——也就是被设定的能让参与人获得较高支付的策略——扩张得更快，而较劣势的策略将逐渐被淘汰。在生物界，这种优胜劣汰的机制通过基因遗传来进行。在商业和社会的策略博弈里，这套机制通常更可能是社会性或文化性的，例如观察和模仿、教育和学习、较成功的投资容易获得更多资金，等等。

进化过程的动态是研究的目标所在。进化过程是否会收敛到一个进化稳定状态？最终只有一个策略大获全胜，还是有若干不同策略共存？有趣的是，许多博弈中进化稳定极限（evolutionary stable limit）与理性算计的参与人达成的均衡状态是一样的。因此进化分析方法为我们采取均衡分析方法提供了有力的后盾。

进化博弈把很多生物学概念引入了博弈论；同样，博弈论也深刻地影响了生物学。生物学家已经发现与其他动物之间的策略互动是动物行为的重要部分。同种动物之间彼此争夺地盘和配偶，异种动物之间则交织成食物链，这些博弈的支付反过来影响着物种的繁衍和进化。正如博弈论在分析选择和动态时受益于从生物学中引入的一些思想一样，生物学也得益于从博弈论引进的策略与支付的思想，这有助于分析动物之间的互动关系。这正是协同和共生的好例子。本书将在

第 13 章对进化博弈的研究提供一个介绍。

2.3.7 观察与实验

整个第 3 节关注的是如何思考博弈或如何分析策略互动。这正是博弈论的精髓。本书将以实例代替严谨的数学或定理推导，介绍初步的博弈理论，虽然初步但理论的本质不变。所有理论必须从两个方面联系现实：现实应有助于建构理论，并用于检验理论是否正确。

我们用两种方法来寻找策略互动的现实基础：(1) 通过对其自然发生时的观察；(2) 设计出一些实验以追踪某些条件带来的效果。我们将在适当的时候通过一些例子对这两种方式进行说明。

许多学者已在课堂上或特别的实验环境中，通过自愿受试参与人来研究策略互动——也就是参与人行为以及博弈结果，其中包括拍卖、讨价还价、囚徒困境以及许多其他类型的博弈，而得到的结果也各有不同。有些理论得到了证实，比如在涉及买卖行为的实验博弈中，买卖双方通常很快可以达到经济理论上的均衡点。有些实验结果则与理论推测出入甚大，例如囚徒困境与讨价还价博弈的实验结果所导致的合作行为，比我们假设参与人自利、追求最大利益时所预期的更多，而拍卖市场实验则存在投标价过高的现象。

我们将在后续章节的若干地方介绍一些由观察和实验所获得的知识，讨论它们与理论之间的联系，并思考如何通过这些知识对理论进行重新诠释、扩展和修正。

2.4 博弈论的用处

本书开篇即已宣称博弈无处不在——无论是个人生活与工作，还是社会、政治、经济的运行；无论是体育赛事，还是其他正式活动；无论是战争中，还是和平时。因此博弈论确实值得系统地学习。不过那只是泛泛之谈，若读者能更清楚地了解博弈论的实际用处，那么学习起来方向就更明确。我们在此列举出三种用处。

用处之一是解释。许多事件及其结果促使我们探询其发生的原因。在存在许多目标不同的决策者进行互动的情况下，博弈论为理解其局势提供了钥匙。举例来说，商战中的割喉式竞争（cutthroat competition）常常是对立双方陷入囚徒困境的结果。本书在适当的地方将提供一些实例，用博弈论分析理解这些事件的发展过程和成因。这中间包括从博弈论角度对古巴导弹危机进行的案例研究。

博弈论的后两个用处是从第一个用处自然地派生出来的。用处之二是预测（prediction）。在观察多名决策者策略互动时，可用博弈预测他们将采取的行动以及结果。当然，在特定背景下的推测还有赖于细节信息，而我们将通过对应用中

的类型广泛的博弈进行分析，让读者培养这种预测能力。

用处之三是提出建议或找出解决的办法：我们可以辅助参与人，告诉他们哪些策略可能获得良好的结果，哪些可能带来糟糕的结果。这类工作仍然是要依实际状况而定的，但我们可以传授给读者一些普遍的原则与技巧，并教他们如何将这些原则和技巧运用到普遍的情形之中。比如第 7 章和第 8 章会说明如何混合采取不同的行动，第 10 章讨论如何使你的保证、威胁和承诺更令人信服，第 11 章将讨论克服囚徒困境的可选方式。

当然，理论本身并不能完美地发挥这三种功能。要解释一个结果，人们必须先正确理解参与人的行为和动机。如前所述，博弈论大多时候都采用一套特殊的方法，即个体参与人理性选择和互动均衡的框架。现实中的参与人与互动性质不见得完全符合这种研究框架。但是，布丁的味道只有自己吃了才知道，理论的好坏唯有自己试了之后才明白。博弈理论分析极大地深化了我们对诸多现象的理解，读者阅读本书后应能相信这一点。而今研究仍在深入，理论仍在改良，本书将给你打下基础，当新理论问世时，让你能更轻松地把握它们并从中受益。

在解释历史事件时，我们常常可以通过历史资料更清楚地了解参与人的动机与行为。而在试图预测或提出建议的时候，则需要猜测参与人行为背后的动机、可能获得的信息或是有限的信息，甚至必须猜测到底有哪些参与人。更重要的是，若其他参与人缺乏算计能力，或其行踪根本就是随机化的，无迹可寻，那么基于博弈理论分析而得出的建议可能就是错误的，因为运用博弈理论进行分析时是假定参与人是理性的、极大化个人利益者而提出建议的。随着越来越多的参与人认识到策略互动的重要，越来越多地使用策略思维或向专家请教，这种错误出现的几率会日益降低，但风险仍然存在（即便如此，博弈论框架带来的系统性分析可以纠正思考策略互动时的逻辑错误，有助于将这些错误发生的可能性降至最低）。此外，博弈论也可将许多不确定性和不完全信息（包括策略的可能性与对手的理性）纳入考虑。我们在后续章节中将提及一些相关的例子。

2.5　后续章节的结构

本章介绍了分析现实博弈问题时所做的思考，要理解或预测任何一个博弈的结果，都必须更深入细致地把握这些思想。本章也介绍了一些对这样的分析非常有用的基本概念。但是一口吃不成胖子，试图一次性掌握所有概念只会带来更多的混淆，从而对每个概念都不甚了解。因此以后我们将一次分析一个概念，以此逐步构建这座理论大厦，并发展分析这些概念的技巧。

在前几章中，从第 3 章到第 8 章，我们将建构并阐释最重要的一些概念和技巧。我们将在第 3 章讨论纯序贯行动博弈并介绍用于分析和求解此类博弈的技术——博弈树和逆向推理。在第 4 章和第 5 章，我们将回到同时行动博弈并提出另一套相应的概念——支付表（payoff tables）、占优（dominance）和纳什均衡（Nash equilibrium）。这两章均关注参与人使用纯策略（pure strategies）的博弈。

第 4 章将参与人严格限定于有限的纯策略集，第 5 章则分析具有连续变量的策略。第 5 章还会讨论更深入的概念问题以及纳什均衡的一个突出可能，即可理性化（rationalizability）。在第 6 章，我们将展现博弈可以综合利用第 3 章至第 5 章中介绍的方法加以研究。在第 7 章，我们回到同时行动零和博弈，这些博弈要求使用随机的或混合的策略。第 8 章将探讨非零和博弈中混合策略的角色，并提出混合策略的一种一般理论。

第 3 章到第 8 章提出的思想和技巧是最基础的部分：（1）用于分析序贯行动博弈的正确的向前看的推理方法（correct forward-looking reasoning）；（2）同时行动博弈的均衡策略，包括纯策略与混合策略。掌握了这些概念和工具，我们就可以把它们运用到第 9 章至第 13 章广泛的博弈和策略的分析中去。

第 9 章讨论参与人受制于不确定性或参与人具有非对称信息的博弈。我们将考察应对风险的策略，甚至是策略性地利用风险。我们也将学习用于操纵和诱导信息的信号传递和信息甄别等重要策略。在第 10 章，我们考虑参与人操纵博弈规则的策略的同时，继续考察参与人操纵规则的作用，亦即先发制人采取策略行动。这样的行动有三种，分别是保证、威胁和承诺。对每一种情况，置信度（credibility）是行动成功的关键，我们会提出一些让这些行动更加可信的方法。

在第 11 章，我们转而研究最广为人知的博弈——囚徒困境。我们将讨论合作是否能够维持以及如何维持，尤其是在重复博弈或长期持续的关系中（然后，在第 12 章我们转向众多而不是少数几个参与人进行策略互动的情况，讨论集体行动博弈）。每个参与人的行动都会对其他参与人造成正面或负面的影响，而结果通常并非对整个社会有利。我们将说明这种结果的本质，并讨论一些可以导致较好结果的简单策略。

所有上述理论和应用均基于这样的假设，即博弈的参与人完全了解博弈的性质，并实施精心算计好的最符合其利益的策略。这种完全理性的假设有时要求太多的信息和计算能力，无法精确反映人们行动时的实际状况。因此第 13 章将从完全不同的角度考察博弈。在此，参与人并不精心算计也不追求最优策略。相反，每个参与人都会选择某一特定的策略，好像基因遗传决定似的。

不同的人在不同的博弈中会有不同的策略，这样的一群参与人相遇在一起并采取各自的策略时，哪个策略会表现得更好？假如较好的策略不断被继承或模仿而成为众多参与人的最优策略，那么这个群体的结构会呈现什么模样？结论是这种进化动态过程通常恰恰垂青于那些会被理性的利益极大化者所采用的策略。因此，进化博弈理论间接支持了我们在先前章节所讨论的最优策略选择与均衡。

在最后一部分——第 14 章到第 18 章，我们将提出策略互动情形的具体应用。在这里我们必须用到来自先前章节的概念和方法。我们从威胁的一个特别有趣的动态版本（即广为人知的边缘政策）开始。我们在第 14 章说明边缘政策的性质并把它运用到对 1962 年古巴导弹危机的讨论中。第 15 章论述的是投票和选举。我们将考察现有的不同的投票规则以及一些可能发生的悖论性（paradoxical）结果。另外，我们将同时研究各种选举中投票者和候选人可行的策略行为。

第 16 章到第 18 章将讨论有价经济资源的配置机制：第 16 章研究拍卖机制，第 17 章考察谈判，第 18 章讨论市场。对拍卖的讨论重点关注招投标双方制定最

优策略过程中，信息和对风险的态度所起的作用。我们也趁机将理论运用到最新类型的拍卖——网上拍卖。第 17 章从合作和非合作角度考察谈判。最后，第 18 章将基于谈判理论中的一些概念和关于核的某种理论（some theory of the core）来考虑市场交换博弈。

所有这些章节提供了大量的学习内容，兴趣不同的读者和教师们应如何各取所需？第 3 章、第 4 章、第 6 章和第 7 章包含了全书的核心理论思想。第 10 章、第 11 章对一般的博弈与策略课程颇为重要。此外还有诸多可选的内容。第 5 章、第 8 章考虑了一些更高深的专题，在理论和数学方面探讨得更深入，这对理工背景较强的学生较有吸引力，社会或人文学科等数理背景较弱的读者则可以忽略这些章节，并不会影响对后续篇章的理解。第 9 章针对一个重要的主题，即大多数现实博弈都涉及不完全或不对称信息，参与人对信息的操纵就成为了许多策略博弈的重要方面，然而用于分析信息博弈的概念与技巧从来就比较复杂。因此读者和教师可以仅仅选择实例，了解信号传递和信息甄别的基本思想而省略其他细节。然而由于这个主题的重要性，我们将这一章放在了第三部分的最先位置。第 12 章、第 13 章均考察有大量参与人的博弈。第 12 章的重点是社会互动；第 13 章的重点则是进化生物学。第 13 章的主题会让那些喜欢生物学的人兴趣盎然，但同样的主题也正在社会科学中浮现，具有社科背景的学生即便要跳过细节也最好要把握这些思想的要点。第 14 章、第 15 章讨论的主题来自政治科学——分别是外交和选举。第 16 章到第 18 章的主题来自经济学——拍卖、议价和市场。针对不同的受众，讲授该课程时应有不同的侧重点，并依据不同的侧重点适当展开。

无论你来自数学、生物学、经济学、政治学或其他学科，还是来自历史学和社会学等领域，策略博弈的理论和例子都将启发你的智慧。无论你是传授还是学习策略博弈，都希望你从中享受到乐趣。

2.6 总　结

策略博弈与个人决策行为的不同之处，在于参与人之间显著的策略互动。博弈可根据行动时序、参与人利益是否彼此冲突、博弈进行的次数、可获取信息的数量、规则类型以及协调行动的可能性等等划分成各种类型。

了解博弈结构中的术语对于分析颇为重要。参与人选择不同的策略，导致不同的博弈结果，获得不同的支付。支付涵盖了对参与人有意义的所有考量，当结果是随机的或包含风险的，则以期望值来计算支付。

博弈论假设所有参与人都具有理性或一致性的行为，并且都了解博弈中必须遵守的所有规则。当所有参与人都采取是对他人策略的最优反应的策略时，均衡就会产生。有些类别的博弈允许从经验中进行学习，人们可以通过这些博弈对通向均衡的动态行为进行研究。对现实博弈状态下的行为进行研究，可以为理论的有效性提供更多的信息。

在不同的环境下，博弈论可以用于解释、预测或提出建议。尽管在这几方面都未能至臻完美，但理论仍在持续发展；而策略互动和策略思维的重要性也越来越广泛地为世人了解和接受。

关键术语

非对称信息（asymmetric information）
合作博弈（cooperative game）
决策（decision）
均衡（equilibrium）
进化博弈（evolutionary game）
期望支付（expected payoff）
博弈（game）
不完美信息（imperfect information）
不完全信息（incomplete information）
非合作博弈（noncooperative game）
支付（payoff）
理性行为（rational behavior）
甄别（screening）
甄别机制（screening devices）
序贯行动（sequential moves）
信号（signal）
信号传递（signaling）
同时行动（simultaneous moves）
策略博弈（strategic game）
策略（strategies）

习　题

1. 下列情景何为博弈，何为决策？你依据什么特性进行分类？
   (a) 一群杂货店老板选择要购买何种口味的酸乳酪。
   (b) 两个少女选择在舞会时要穿的衣服。
   (c) 一位大学生考虑要接受何种研究生教育。
   (d) 微软公司和网景公司选择它们的互联网浏览器的价格。
   (e) 一位州长候选人选择竞选搭档。
2. 考虑下列博弈。每一种情况下，请根据课本中给出的六个方面对博弈进行分类，说明博弈属于下列哪一类：同时行动或序贯行动、博弈是否为零和、博弈是否重复、是否不完美或不完全（非对称）信息、规则是否固定、合作协议是否可能。如果你没有足够的信息按照某一方面对博弈进行分类，请说明理由。
   (a) 剪刀—石头—布：数到三时，每位参与人出拳。石头赢剪刀，剪刀赢布，布赢石头。
   (b) 唱名投票：投票人在被叫到其名字时以口头方式投票。两位候选人中高票者当选。
   (c) 密封投标拍卖：投标者将标价写好封于信封中。出价最高者可依其标价获得标的物——一瓶红酒。
3. “面对一场所有参与人皆可获得一定利益的博弈，相较于一场某参与人可以独享全部好处的博弈，该参与人更愿意选择后者”这一论述是否正确？请简单

说明理由。

4. 你和对手正进行着这样一场有三种可能结果的博弈：你赢、你对手赢（你输）、两人平手。如果你赢则你得到支付 50 元，若平手则得到支付 20 元，如果你输则得到零支付。计算下面三种情况下你的预期支付：

   (a) 你有 50%的概率平对手，10%的概率赢对手（从而你输的概率是 40%）。

   (b) 胜负概率各为 50%，不可能打成平手。

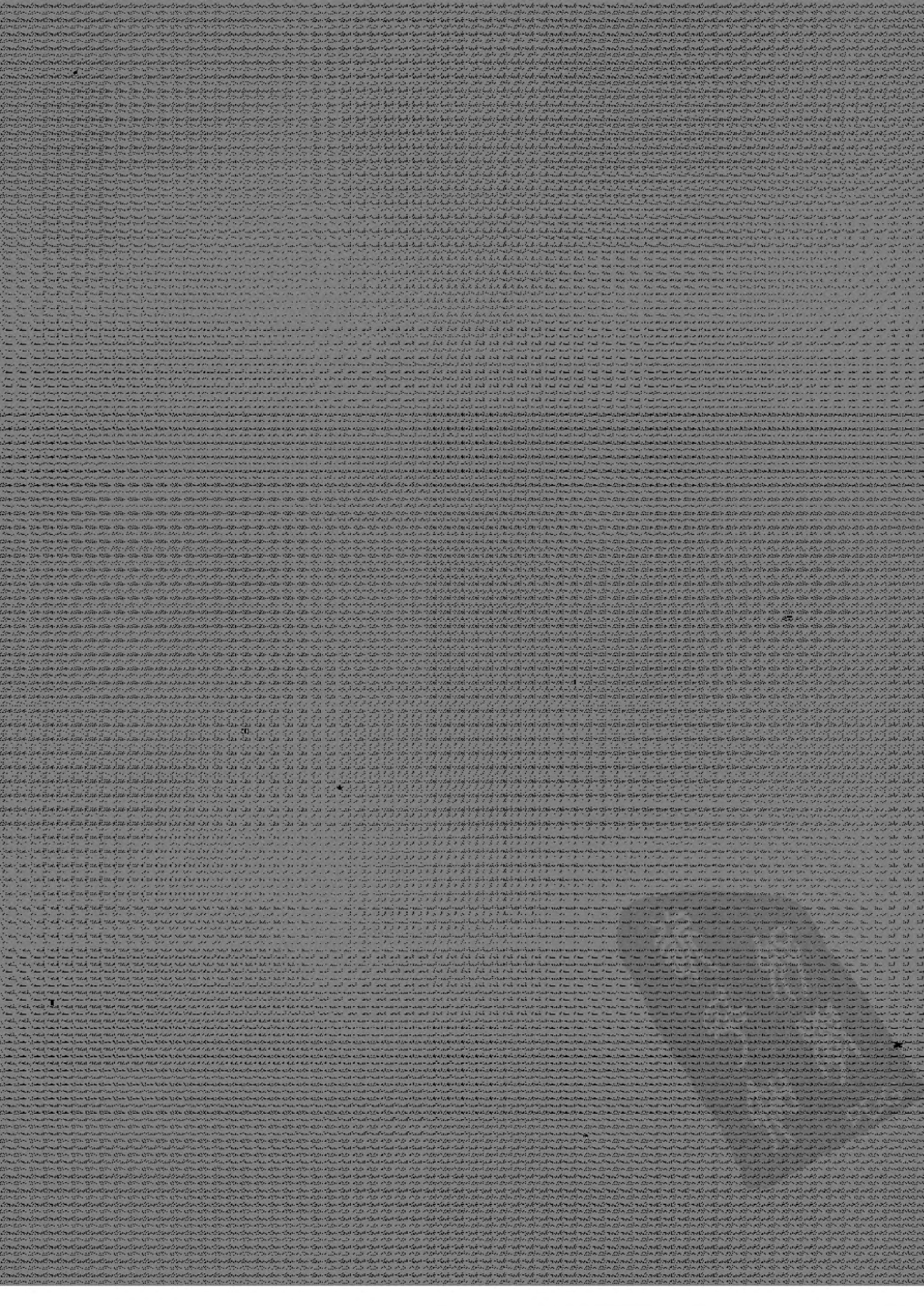
   (c) 你有 80%的概率会输，赢或平手概率各为 10%。

5. 解释博弈论作为预测工具和作为建议工具时有何不同。在哪些现实情景中这两者显得特别重要？

【注释】

[1] James M. McPherson, "American Victory, American Defeat," in *Why the Confederacy Lost*, Ed. Gabor S. Boritt, New York: Oxford University Press, 1993, p. 19.

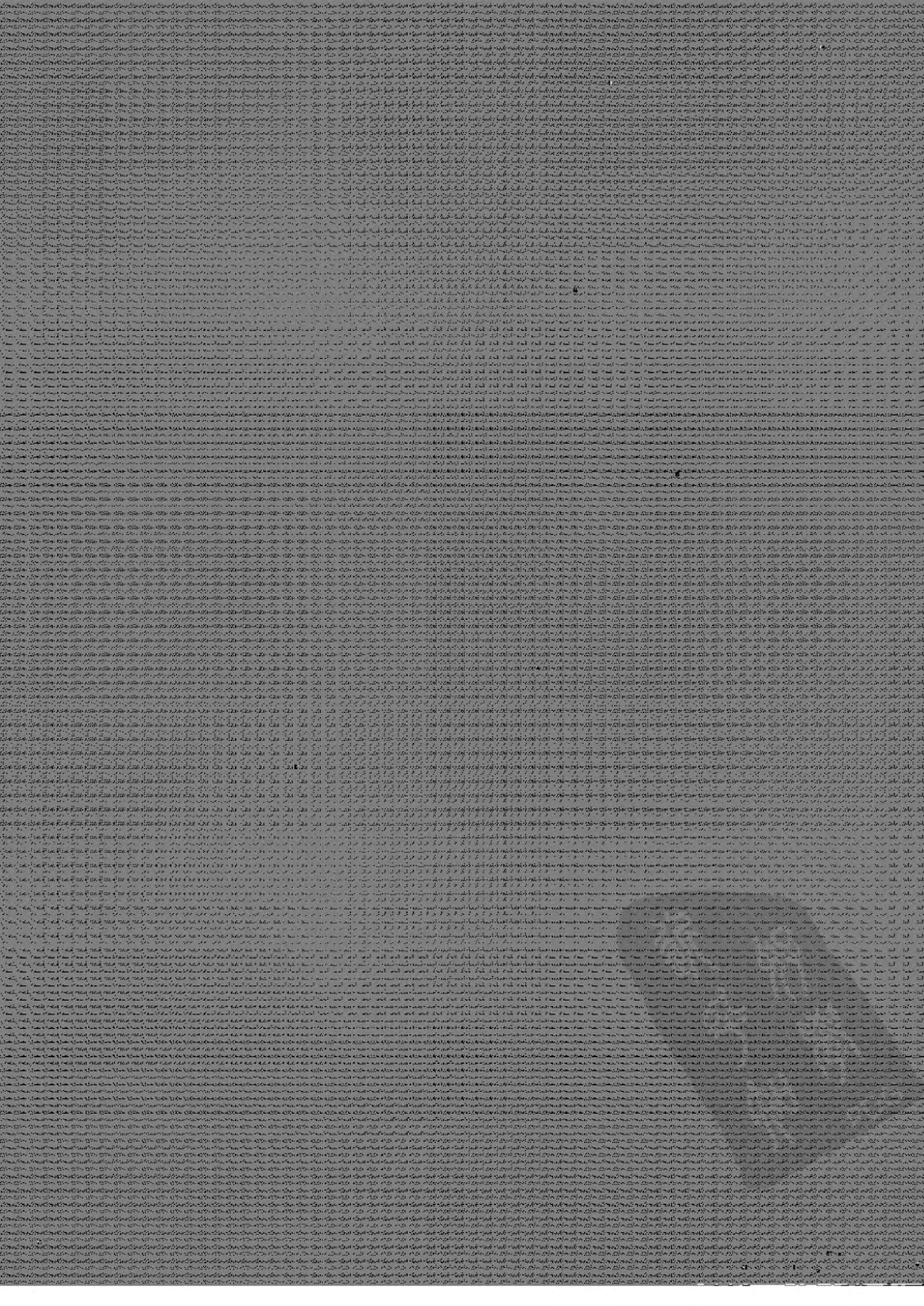
[2] 若要进一步了解洛克菲勒取得权力所使用的方法，可参考 Ron Chernow, *Titan*, New York: Random House, 1998。



第二部分

概念与技巧





第3章　序贯行动博弈

序贯行动博弈涉及这样一类策略性互动的情形，其中参与人的行动存在着严格的先后顺序。参与人轮流选择其行动，并且他们都知道在他们完成其行动选择之前有哪些参与人进行了行动选择。为了在这种博弈中获得尽可能好的结果，参与人一定要运用特定类型的互动性思维。每一个参与人都要考虑到：倘若我自己选择了这样的行动，我的对手将会有什么样的反应？无论何时完成了行动选择，参与人都需要料想他们当前的行动会如何影响未来的行动，包括他们对手的行动及自己的行动。于是，参与人在计算未来结果的基础上决定他们当前的行动选择。

大多数进行的博弈都包含了序贯及同时行动的方面。但是，倘若首先将它们分解为两种单纯的情形加以研究，则分析方法及有关概念会更加易于理解。因此，在本章我们将单独研究纯序贯行动博弈。第4章和第5章将着手研究纯同时行动博弈，而第6章、第7章及第8章中的有关部分将表明如何在更加现实的混合情况下把两种类型的研究结合起来。这里给出的研究可被用于那些含有序贯性决策制定过程的博弈分析。序贯行动博弈分析还对在什么时刻参与人先行动会具有优势和什么时刻后行动更好提供了一些信息。这样，参与人就可以设计一些方法（称为策略行动）以操纵博弈的顺序从而获得优势。对这种问题的分析是第10章讨论的焦点。

3.1 博弈树

我们从介绍一种展示和分析序贯行动博弈的几何技术开始，这种几何方法被称为**博弈树**（game tree）。这种树形图被称为博弈的**扩展式**（extensive form），它将我们在第 2 章中引入的有关博弈的所有组成要素如参与人、行动及支付等都一一标示出来。你或许在其他课程中曾遇到过**决策树**（decision tree）的概念。这类树形图表达的是单个决策者在一个中性环境中连续不断的决策点或决策结。决策树也包括了对应于从每个决策结发出的可选择行动的枝。博弈树正是博弈中所有参与人决策树的合并。这种树形图表明了所有参与人可选择的所有可能行动，并给出了博弈的所有可能结果。

3.1.1 决策结、枝及路径

图 3—1 给出了一个特定序贯行动博弈的树形图。我们并不给出这个博弈的故事背景，因为我们打算忽略掉具体的环境细节而将注意力集中在一般性概念上。这个博弈有四个参与人：安妮、鲍勃、切蕾丝以及笫勃。博弈的规则给予安妮首先行动的安排；在最左端的**结点**（node）上表达出这种含义，该结点被称为博弈树的**初始结**（initial node）或**根**（root）。在这个也可称为**行动结**（action node）或**决策结**（decision node）的结点上，安妮有两个可选择的行动。安妮可选择的行动用“停止”和“继续干”加以标明（记住，这些标记是抽象的，并没有特定的意义），并且用从初始结发出的**枝**（branch）来加以表示。

如果安妮选择“停止”，则轮到鲍勃开始行动了。在他的行动结上，他有三个可选择的行动，分别用 1、2、3 标出来。倘若安妮选择“继续干”，则由切蕾丝第二个行动，她有“冒险”和“安全”两个选择。接着可类似地进行其他结点和枝的说明；但是，我们不打算将它们都一一用文字表述出来，只是将你的注意力引导到一些重要的特征上去。

如果安妮选择“停止”且鲍勃随后选择“1”，又轮到安妮选择，现在她有新的可选择行动，即“上”和“下”。在实际的序贯行动博弈中，一个参与人有几次行动选择的机会并且在不同的时候有不同的可选择行动是很常见的。譬如，在国际象棋比赛中，两个参与人交替地行动，每一次行动都改变了布局且因此改变了不同轮次的可选择行动。

3.1.2 不确定性与“自然的行动”

如果安妮选择“继续干”，则切蕾丝选择“冒险”，随机性就出现了——抛一

枚硬币且博弈结果由硬币出现的是“正面”还是“反面”来决定。博弈的这种特征在树形图中是通过引入一个被称为“自然”（nature）的外部参与人来处理的。随机性事件被假定为由被称为自然的参与人来控制，正如图中所示，它选择每一个枝的概率是50%。在这里，由随机性事件的类型即抛硬币决定了一个固定不变的概率，但在其他的场合它可能是不同的，譬如，扔骰子时会出现六个可能的结果，各以约16.67%的概率出现。采用引入“自然”这个参与人的方法，我们得以在博弈中引入不确定性，也使我们可以将某些超出任何实际参与人的控制之外的事件纳入考虑范围。

你可以通过下述由枝形成的链在博弈树中构造出许多不同的路径。在图3—1中，每条路径都经过有限次行动将你带到博弈的某个终点处。博弈不一定有终点，某些博弈在原则上可以无限进行下去。但是，我们所考虑的大多数应用问题都是有限博弈。

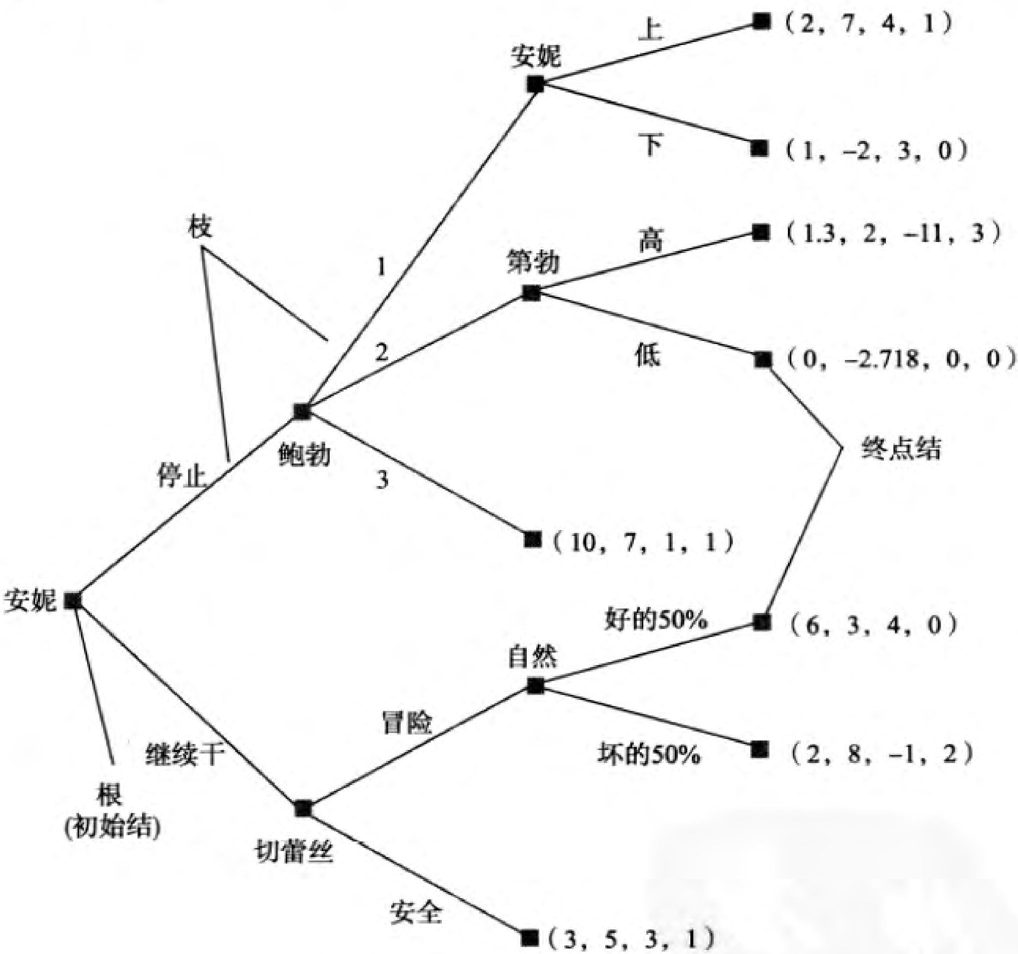


图3—1　博弈树的一种说明

3.1.3 结果与支付

沿每一条路径都存在一个最后的结点，即**终点结**（terminal node）。在该结点，所有参与人都不再进行行动选择（注意，终点结因此与行动结是不同的）。

我们用参与人的支付来表示这些特定行动序列的结果。就这里四个参与人的情形而言，我们按照这样的顺序（安妮、鲍勃、切蕾丝、第勃）来写出支付。重要的是要将支付与对应的参与人联系起来。通常的办法就是以向量形式按照参与人行动的先后顺序写出支付。但是，这种方法有时也会产生歧义。在我们的例子里，不太明确的是到底第二个行动的是鲍勃呢还是切蕾丝。因此，我们采用词典排序法（alphabetical order）。[①] 当你用博弈树来分析一个博弈时，你可用任何你认为方便的方法来标记，但你一定要对读者说明。

支付是用数字来表示的，并且一般说来，对于每一个参与人，一个较高的数字就意味着一种较好的结果。所以，对于安妮来说，图 3—1 中最底端的路径结果（支付 3）要比最上端的路径结果（支付 2）好一些。但是，在不同参与人之间并不存在这种可比性。因此，在最上端的路径的末端，鲍勃（支付 7）要比安妮（支付 2）好一些这种说法不一定有什么意义。譬如，如果支付是用金钱来衡量的，这样一种在不同个人之间的比较可能就有意义。

参与人在可以挑选的各种各样行动之间进行决策的时候会使用有关支付的信息。引入一个随机性事件（自然的选择）就意味着参与人需要决定他们在自然行动的时候会在平均意义上得到什么。例如，如果安妮在博弈的第一个行动上选择了“继续干”的话，切蕾丝就会随后选择“冒险”，导致像抛硬币那样由自然在“好的”或者“坏的”之间进行“选择”。此时，安妮能够预料到有一半的几率获得支付 6 和另外一半的几率获得支付 2，或者一种统计平均或者说是预期支付 $4=(0.5\times6)+(0.5\times2)$。

3.1.4 策　　略

最后，我们运用图 3—1 中的博弈树来说明策略的概念。参与人在结点上所选择的单个行动被称为**一步**（move）。但是，参与人可以制定这样一个计划，它将他们预期在博弈过程中将可能出现的各种各样场合下可能会选择的行动事前规定好。这种行动计划就被称为一个策略。

在这个博弈树中，鲍勃、切蕾丝和第勃每人最多有一次行动机会，譬如，切蕾丝仅当安妮首先选择“继续干”时才有机会行动。对于他们来说，一个行动与一个策略之间是没有区别的。我们可以通过对特定情形下的行动选择加以说明来刻画策略。因此，鲍勃的一个策略可以写成：“如果安妮选择了‘停止’，就选择‘1’。”但是，安妮却有两个机会进行选择，故其策略需要进一步完全地刻画。她的一个策略便是：“选择‘停止’，且当鲍勃选择‘1’后选择‘下’。”

在像国际象棋这样十分复杂的博弈中，有很长的行动序列，而且每一次行动选择时都有许多可选择的行动，描述其策略过于复杂。我们将在本章稍后的章节

① 原书中还用不同的颜色区分安妮、鲍勃、切蕾丝、第勃，并且在图中通过不同颜色的数值区分不同的支付。我们在本译本中不通过颜色来进行这样的区分，而是通过增加文字说明来达到这一目的。在本译本中对其他图形也加以类似处理，后面就不再一一说明。——译者注

再来考虑这方面的问题。但是，除了一点特别需要注意之外，构造策略的一般性原则是十分简单的。倘若安妮在开始选择的是“继续干”，她就不会有机会做第二次选择。在一个选择了“继续干”的策略中，需要规定如果处于第二次行动的结点上这种假设情形里她应该如何行动吗？你的直觉可能回答“不”，但规范的博弈理论的回答则是肯定的。之所以如此，有两点理由。

首先，安妮在开始做出“继续干”的选择可能会受到她对她在第二次行动中应做何选择这种考虑的影响，倘若她在开始时做出“停止”的相反选择的话。譬如，如果她选择“停止”，鲍勃就会随后选择“1”；于是安妮就有第二次行动的机会，此时其最好选择是“上”，使她得到支付 2。倘若她开始选的是“继续干”，切蕾丝会选择“安全”（她从选“安全”中得到的支付是 3，好于从选择“冒险”中得到的预期支付 1.5)，并且这个结果会给安妮带来支付 3。为使这个思路过程得以更加清楚地表达，我们将安妮的策略表述为：“开始选择‘继续干’，如果出现第二次行动机会，则选择‘上’。”

进行这种看起来十分繁琐的策略表述的第二个理由与均衡的稳定性有关。当谈及稳定性时，我们要问倘若参与人的选择受到微小扰动，将会发生什么事情。一种这类扰动就是参与人的选择出现小小的错误。譬如，如果是通过按键的方式来完成行动选择，安妮本来打算按的是“继续干”键，但存在一个小的概率使得她的手发抖，从而使她实际上按了相反的“停止”键。这里，重要的是当安妮因鲍勃选了“1”而发现了自己的错误之后，且又轮到由安妮来做出行动选择时，应如何规定安妮进行行动选择所应遵循的规则。更高深水平的博弈理论要求进行这样的稳定性分析。我们坚持要求你用这样一种完整的行动计划来刻画策略，就是打算要你从一开始就为此做好准备。

3.1.5 博弈树的构造

现在，我们来将图 3—1 中的博弈树所表达出的一般性概念加以总结。

博弈树由结点和树枝组成。不同的结点由树枝连接起来，并且有两种类型的结点。第一类结点被称为决策结。与每一个决策结相联系的是一个参与人，他在结点处进行行动选择；每一个博弈树都有一个决策结作为博弈的初始结，是博弈的起点。第二类结点被称为终点结。与每一个终点结相联系的是一个结果集合，这个结果集合是参与人在博弈中所获得的博弈结果；如果博弈沿着导致这些特定终点结的树枝进行，则这些结果就是每个参与人所得到的支付。

博弈树的枝表示从任意决策结出发能够选取的可能行动。从博弈树上一个决策结发出的每一个枝或者走向另外的决策结——一般是属于另外参与人的，或者指向一个终点结。博弈树必须把每个结点上参与人所有可能做出的选择都考虑在内；所以，某些博弈树包括了一些枝，这些枝对应的选择是“什么都不干”(do nothing)。每一个决策结至少发出一个枝，但发出枝的数量是没有上限的。然而，每个决策结只能有一个枝指向它。

通常把博弈树从左向右画出来。但是，只要方便，可以从任何方向开始画博

弈树：从下向上、从侧面开始、自上而下，或者甚至从一个中心点辐射状地向外画。博弈树这种说法是一种隐喻，重要的是随着在博弈树的各结点上做出决策，博弈是沿着树枝接连有序进行的这样一种思想。

3.2 用博弈树求解博弈

抽烟还是不抽烟？大多数人都可能曾面临过这样的决定。我们将以这一非常简单的例子来说明博弈树在寻找序贯行动博弈均衡结果过程中的运用。这个问题以及许多其他类似的单个参与人策略选择问题都可归入博弈问题而加以解决，如果我们认识到未来的选择实际上是由另一不同的参与人做出的话。这一参与人就是人们在未来的自己，他受到不同的影响且对博弈的理想结果有着不同的观点。

来看看一位叫做卡门的少女，她在决定是否要抽烟。首先，她必须要决定是否试着抽一下。如果她试了一下，她还要进一步决定是否继续抽。我们用图3—2中的博弈树中的一个简单决策来说明这个例子。在枝上标示着卡门的可选择行动，但我们还需要说明支付是什么。将根本就不抽烟的结果作为参考标准，且令其支付为0。在这里，数字0并无特别的含义，对于结果的比较，从而对于卡门的决策来说，重要的是最后所获得支付之间的相对大小。假设卡门最喜欢的结果是她试着抽一下但不会继续抽下去。理由可以是她很喜欢自己去经历不同的事情或者当她在未来试着阻止她的孩子抽烟时，她能够更加自信地说："我抽过，但那不是件好事。"令这一结果的支付为1。另一结果——她试着抽了一下且继续抽下去——是最糟糕的了。即使先把长期健康风险放在一边不去考虑，也存在着当前的问题——她的头发和衣服会染上烟味，并且她的朋友们也会疏远她。令这个结果的支付为－1。卡门的最佳选择看来就明了了——她应试着抽一下但不应持续抽下去。

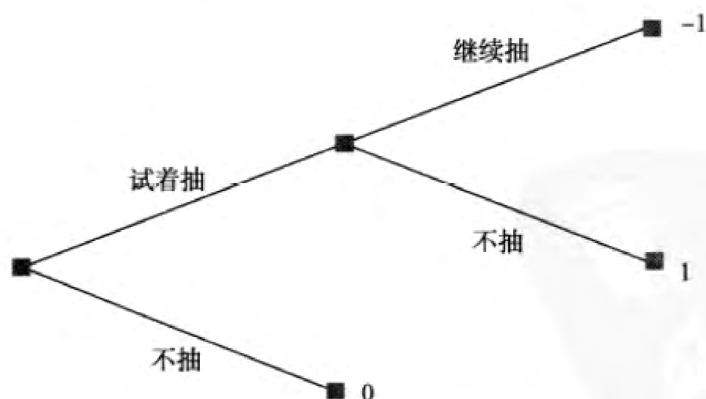


图3—2 抽烟决策

然而，这种分析却忽略了烟瘾的问题。一旦卡门试着抽了一段时间烟，她会变成另外一个人，有着不同的嗜好，从而有不同的支付。决定是否继续抽烟的并不是"今天的卡门"——她对抽烟之后果的评价由图3—2给出，而是一个不同

于今天的卡门的“未来的卡门”——其对不同后果的评价是不同于此时的。当今天的卡门做出选择时，她必须预期到其结果，并将这一结果纳入其基于当前偏好的当前决策的考虑范围。换句话说，与抽烟有关的决策问题并不是在第 2 章中所说明的那种意义上的决策——单个人在一个中性环境中做出的决定——而是一种博弈，也是在第 2 章中说明的那种技术意义上的博弈，其中其他参与人就是未来具有不同偏好的卡门自己。当今天的卡门做出决策时，她不得不与未来的自己进行博弈！

我们通过将在两个结点上进行决策的两个参与人加以区别，就将图 3—2 中的决策树转换为图 3—3 所示的博弈树。在初始结上，“今天的卡门”在决定是否试着抽烟。倘若她决定试一下，则增加的“未来的卡门”就进来了，并且要选择是否继续抽。今天的卡门的支付与之前相同。但未来的卡门在继续抽烟中获得舒服的享受，但一旦停止抽，她就会感到相当难受。设未来的卡门从“继续抽”中获得支付 1，从“不抽”中获得支付 −1。

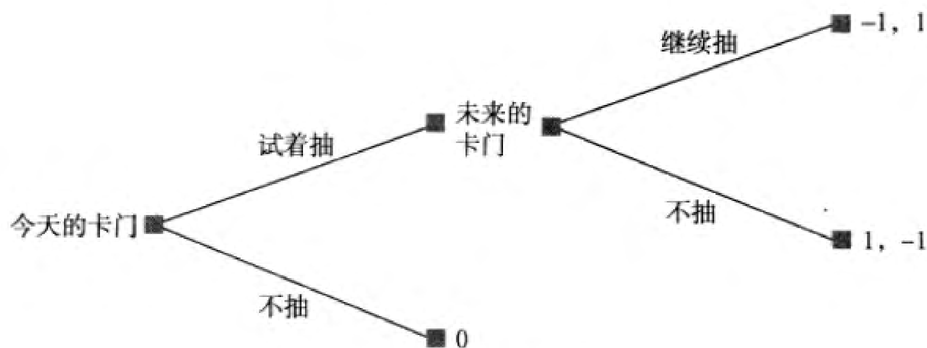


图 3—3　抽烟博弈

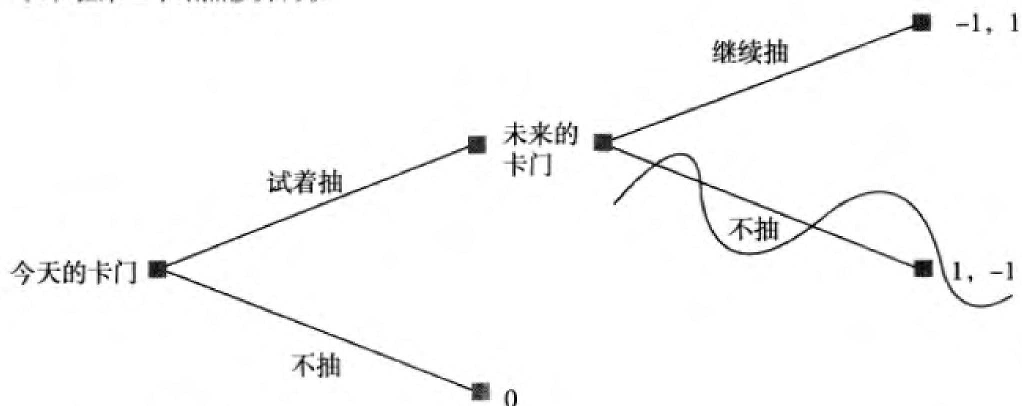
给定已成为瘾君子的未来的卡门之偏好，她将在其决策结上选择“继续抽”。今天的卡门会预想到这一前景，并且将其纳入当前决策考虑之中，从而认识到选择试着抽一会儿必定会导致其“继续抽”。即使在给定当前偏好下今天的卡门并不打算继续抽下去，在未来她也不能贯彻当前偏好的决定，因为这一决策是由一个具有不同偏好的不同的卡门在未来做出的。所以，今天的卡门将预见到“试着抽烟”的选择会导致“继续抽”，并且使其获得在今天看来是 −1 的支付，而“不抽”的选择将使其获得支付 0。因此，她会选择后者。

在图 3—4 中，我们更为规范和直观地表达出这一论证。在图 3—4（a）中，我们将从第二个结点发出的“不抽”树枝“剪去”（prune）。这种“剪去”对应于这样一个事实，即在这个结点上进行行动选择的卡门不会选择与此一树枝相对应的行动，给定其偏好如图中所示情况下一定如此。

剩下的树仍有从第一个结点发出的两个枝，今天的卡门在这个结点进行她的选择；每一条枝现在都指向一个终点结。这种修剪使今天的卡门完全预测到其每个选择的最终后果。在“试着抽”之后接着便是“继续抽”，且带来支付 −1，这个支付是用今天的卡门的偏好来衡量的；而“不抽”会带来支付 0。这样，卡门今天的选择就会是“不抽”而不是“试着抽”了。所以，我们就能够剪去从第一个结点发出的枝“试着抽”（以及其可预见到的后续过程）。这一修剪出现在图

3—4（b）中。现在出现的树是"充分修剪过的"了，从初始结发出的枝只有一条了，并且它指向一个终点结。沿着这唯一留下的穿过整棵树的路径，表明当所有参与人在正确预见到全部未来后果下进行其最优选择时，在博弈中将出现什么。

（a）在第二个结点修剪树枝



（b）充分修剪树枝

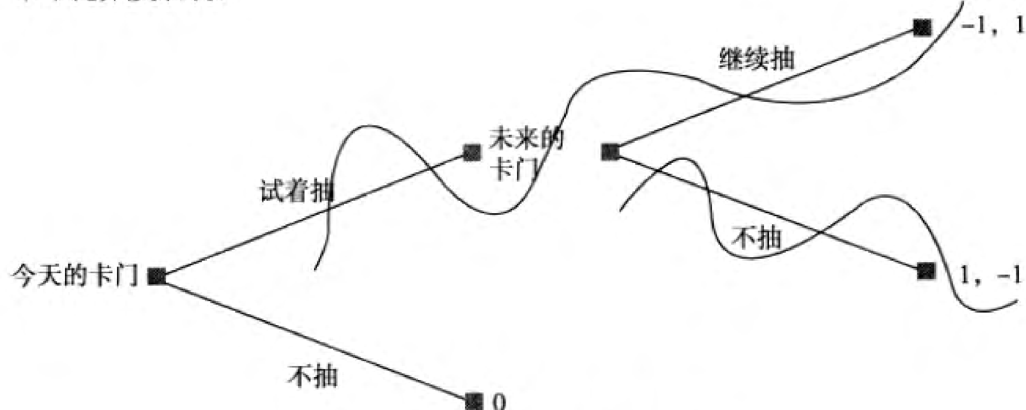


图 3—4　修剪抽烟博弈的树

在图 3—4 中的修枝过程中，我们划掉了未选的树枝。另外一个等价但不同的表示参与人选择的方法是将被选择的树枝标示出来。为此，你可以在这些枝上标示显著的记号或把箭头放在这些枝上，或者用较粗的线来表示它们。所有这些方法都可以用，在图 3—5 中将它们都表示了出来。你可以选择是"修剪"还是用记号表示，但是后者，特别是用箭头的形式有一些优势。首先，它带来清晰的图像。其次，剪去树枝的图像之混乱状态有时无法显示出各个枝被剪去的顺序。例如，在图 3—4（b）中，读者可能会搞混，从而错误地认为在第二个结点上的"继续抽"枝会被首先剪去，紧接着，在第二个结点上衍生出"不抽"枝的第一个结点上的"试着抽"枝会被剪去。最后，也是最重要的，用从初始结到终点结的箭头连续连接起来的方式，最为醒目地显示出了最优选择序列的结果。因此，在随后的这种类型的图形里，我们一般采用箭头而不是修剪树枝的方法。当你画博弈树时，你应该尝试一下两种画法；当你习惯了博弈树时，你就可以按照你的偏好选择其中的任何一种方法。

在博弈树中你如何表达你的思路是无关紧要的，分析的逻辑是相同的，也是

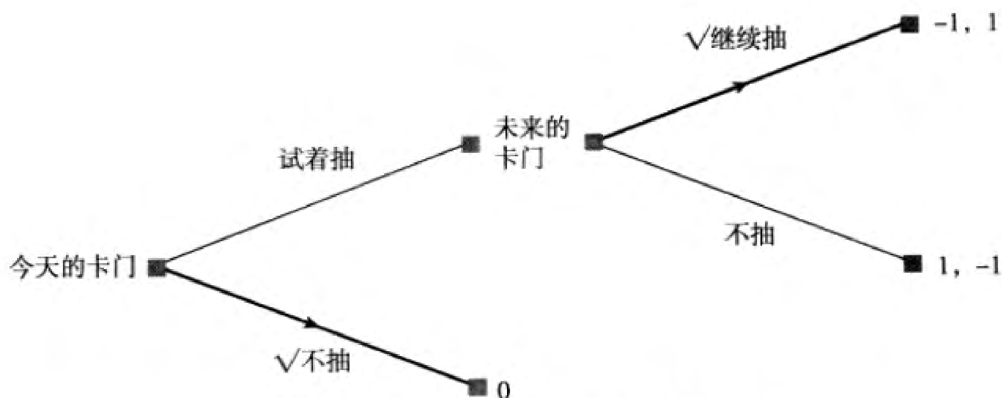


图 3—5　显示抽烟博弈中的树枝选择

重要的。你必须从考察那些直接通向终点结的行动结开始分析。通过比较其在相应终点结上的支付，就可立即找出在这样一个结点上行动的参与人的最优选择。通过采用这种博弈终点性选择去预测前面选择行动的结果，在最后的决策结前面一个结点上的选择就可被决定了。然后，对于再前面的结点也可依同样方式选择，依此类推。沿着博弈树按照这种方法向后进行，你就可以解出整个博弈。

这样一种通过向前看且向后推理来决定序贯行动博弈中行为的方法被称为**逆推**（rollback）。恰如这个名称所表示的那样，当你展开分析的时候，运用逆推要求在一开始就思考在所有的终点结上将会发生什么，并且通过博弈树直到初始结一步一步地"逆推"。由于这种推理要求在每一个时期进行一步反推，该方法也被称为**逆向归纳法**（backward induction）。我们采用"逆推"这一术语是因为它更加简单且已经被更加广泛地采用了，但其他博弈论文献也用逆向归纳法这一较老的术语。切记这两个术语是等价的。

当所有参与人都通过运用逆推分析法选择其最优策略时，我们称这样的策略集合为博弈的**逆推均衡**（rollback equilibrium）；称这种策略所带来的结果为逆推均衡结果（rollback equilibrium outcome）。博弈论预言这种结果是所有参与人都是追求各自的最优化支付的理性算计者时的序贯行动博弈均衡。在本章的稍后部分里，我们要强调这种预言在实践中在多大程度上是合理的。在当前，你应该知道本书中给出来的所有有限序贯行动博弈都至少有一个逆推均衡。事实上，大多数正好只有一个。仅仅在其中参与人从两个或更多的不同行动集中获得相同支付，且因此在它们之间是无差异的例外情形里，博弈将有多于一个的逆推均衡。

在抽烟博弈中，逆推均衡是今天的卡门选择策略"不抽"和未来的卡门选择策略"继续抽"。当今天的卡门执行其最优行动时，瘾君子未来的卡门就根本不会出现且因此实际上没有机会行动。但未来的卡门的影子的存在和当今天的卡门选择"试着抽"且给她一个行动的机会时，她将选择的策略都是博弈的重要组成部分。事实上，它们在决定今天的卡门之最优行动中起着重要的作用。

我们在一个非常简单的例子中引入了博弈树和逆推的概念，其中的解从文字性论证上看是十分显然的。现在，我们来把这些概念逐步地用于更为复杂的场合，其中文字性分析将变得困难一些，而运用博弈树的图形分析将变得更为重要。

3.3 增添更多的参与人

在第 2 节中两个参与人和两种行动的简单框架里提出的技巧可以立即加以扩充。博弈树变得更为复杂一些，有更多的树枝、更多的结点和数字，但基本的概念和逆推的方法仍然不变。在本节，我们来考察一个有三个参与人的博弈，其中每个参与人都有两种选择；稍加变化，该博弈会再次出现于随后的许多章节中。

这三个参与人是爱弥莉、莱娜和塔尼娅，她们都生活在同一条小街上。要求她们每人都要为在小街与主要公路干道的交汇处修建一座花园作出贡献。花园最终的大小规模和壮丽程度依赖于她们将作出多大的贡献。进而，虽然每一个参与人都乐于见到有这样一个花园，且认为规模越大越好，越富丽堂皇越好，但每个人都由于为此作出贡献是有成本的而不是很愿意作出贡献。

假定，如果两个人或所有的三人都作出贡献，则就会有足够的资源进行最初的花草种植和随后的花园维护，花园将因此而非常漂亮和令人心旷神怡；然而，倘若只有一个人或没有人作出贡献，则花园中的花木就十分稀疏且得不到很好的维护，从而令人不快。从每个参与人的角度看，存在着四种完全不同的结果：

- 她不作贡献，而其他两人作贡献（有一个令人心旷神怡的花园，还省去了自己作贡献的成本）
- 她作贡献，且其他的一个或两个人都作贡献（有一个令人心旷神怡的花园，自己付出了成本）
- 她不作贡献，其他人中仅有一人或没有人作贡献（稀疏的花木，自己节省了成本）
- 她作贡献，但其他人中没有人作贡献（稀疏的花木，自己付出成本）

在这些结果中，列于最上面的显然是最好的且列于最下面的显然是最糟糕的。我们用较高的支付数值表示有较高评价的结果，所以，给予最上面的结果以支付 4，并给予最下面一个结果以支付 1（有时也按数字排列顺序给结果赋予支付，从而在四种结果下，1 是最好的而 4 是最差的，且较小的数值表示更加偏好的结果。在阅读时，你要留意作者用的是哪一种方式；在写作时，你要详细说明你用的是哪一种方式）。对于两个中间的结果，存在着某种不确定性。让我们假设每一个参与人都把一个令人心旷神怡的花园看得比自己作的贡献更为重要。于是，列在第二位的结果获得支付 3，且列于第三位的结果得到支付 2。

假定参与人是序贯行动的。爱弥莉首先行动，并要选择是否作出贡献。随后，在观察到爱弥莉的选择之后，莱娜在作贡献和不作贡献之间进行选择。最后，在观察到爱弥莉和莱娜的选择之后，由塔尼娅来进行类似的选择。[1]

图 3—6 给出了这个博弈的博弈树。我们标出了行动结点以方便分析。爱弥

莉在初始结 a 上行动，树枝对应于她的两个选择，即作贡献与不作贡献，分别引向结点 b 和 c。在每一个这样的结点上，莱娜开始行动且在作贡献和不作贡献之间做出选择。她的选择引向结点 d、e、f 和 g，在每一个这样的结点上，就由塔尼娅行动了。她的选择导向 8 个结点结，其中我们按顺序（爱弥莉、莱娜、塔尼娅）给出了支付。[2]例如，如果爱弥莉选作贡献，则莱娜就选不作贡献，且最后由塔尼娅选作贡献，则花园将是令人心旷神怡的，且两个贡献者每个都得到支付 3，而未贡献者获得其最好结果，支付为 4；此时，支付表列为（3，4，3）。

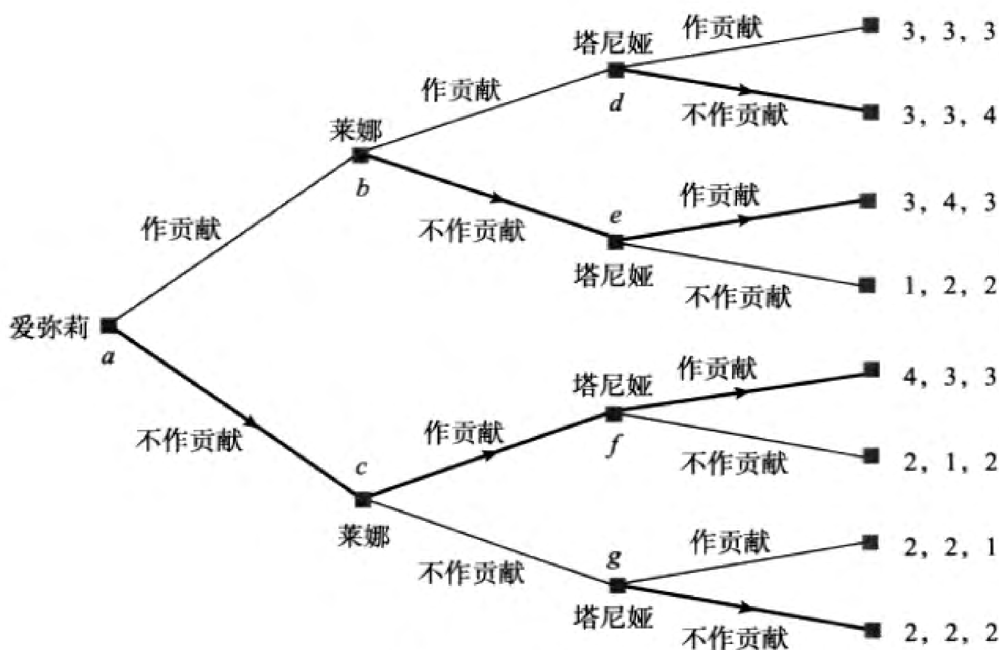


图 3—6　街心花园博弈

为在这个博弈中运用逆推分析，我们从终点结前面的行动结点即 d、e、f 和 g 开始。在每一个这类结点上，都由塔尼娅行动。在结点 d，她面临着的情形是爱弥莉和莱娜都作出贡献。花园一定是令人心旷神怡的了；所以，如果塔尼娅选择不作贡献，她得到其最佳结果 4，相反，倘若她选择作贡献，她就获得次优结果。她在这个结点上所偏好的选择于是就是不作贡献。我们用加粗树枝并在上面画上一个箭头的办法来表示这种偏好，其中每一种都足以表明塔尼娅的选择。在结点 e，爱弥莉作了贡献但莱娜未作贡献，故塔尼娅的贡献对于建一个令人心旷神怡的花园来说是至关重要的。如果她选择作贡献的话，她将得到支付 3；而当她选择不作贡献时得到支付 2。她在结点 e 偏好的选择就是作贡献。你可以类似地验证塔尼娅在其他两个结点上的选择。

现在，我们来对前一阶段即结点 b 和 c 进行逆推分析，这里由莱娜进行选择了。在结点 b，爱弥莉作了贡献。现在，莱娜的推理如下："如果我选择作贡献，将会使博弈走向结点 d，在那里我知道塔尼娅会选择不作贡献，我会得到支付 3（花园将会是令人心旷神怡的，但我会付出成本）。倘若我选择了不作贡献，博弈会走向结点 e，在那里我知道塔尼娅会选择作贡献，而我会得到支付 4（花园将

会是令人心旷神怡的，而我也会省去成本）。所以，我会选不作贡献。”类似的推理表明，在结点 c，莱娜将选择作贡献。

最后，考虑爱弥莉在初始结 a 的选择。她会预见到莱娜和塔尼娅随后的选择。爱弥莉知道，如果她选择作贡献，随后的选择将是莱娜不作贡献和塔尼娅作贡献。在有两个贡献者的情况下，花园将是令人心旷神怡的但爱弥莉会付出成本，所以她的支付为 3。如果爱弥莉选不作贡献，则随后二者的选择都是作贡献，此时会有一个令人心旷神怡的花园，且爱弥莉不会付出成本，爱弥莉的支付将为 4。故她在初始结 a 所偏好的选择为不作贡献。

对于这个街心花园博弈，逆推分析的结果现在可轻易地加以总结出来了。爱弥莉选不作贡献，然后是莱娜选作贡献，最后由塔尼娅选作贡献。这些选择给出了一条通过博弈树的特定**进行路径**（path of play）——沿着从初始结 a 发出的下面的树枝，然后沿随后到达的两个结点 c 和 f 各自发出的上面的树枝。在图 3—6 中，容易看出来进行路径就是从初始结到从上向下数第 5 个终点结由首尾相连的箭头序列构成的连续路径。参与人获得的支付被标记于这个终点结上。

逆推分析是简单而明了的。在这里，我们强调一下其某些特征。首先，注意序贯行动博弈的**均衡进行路径**（equilibrium path of play）未达到大多数的枝和结点。然而，推算到达这些结点后将选择的最佳行动，却是求解最终均衡的一个重要环节。博弈中，优先行动的参与人决定如何行动时，不仅会受他们对自己选择非最优行动的结果的预期的影响，而且还受他们对其他参与人不选择其最优行动的结果的预期的影响。这些建立在偏离均衡结点（与在逆推过程中修剪过的树枝相联系的结点）上的预期，使得参与人在每个结点上选择最优行动。例如，爱弥莉在第一个结点上的最优选择不作贡献受控于这样一种知识，即若她选择作贡献，则莱娜将选不作贡献，接着由塔尼娅选择作贡献；这一序列给爱弥莉带来支付 3，而不是通过在开始就选不作贡献所应获得的 4。

逆推均衡通过确定每一个参与人的最优策略对这种分析的全过程给出了一个完整的陈述。回想一下，之前我们曾提到，一个策略是一个完整的行动计划。爱弥莉首先行动且恰好有两个选择，故她的策略相当简单且当她行动时是同样有效的。但对于莱娜来说，她是第二个开始行动的，并在两个结点之一处行动：若爱弥莉选的是作贡献，则她在一个结点处行动；若爱弥莉选的是不作贡献，则她在另一个结点处行动。莱娜的完整行动计划不得不对她将在每一种情形下应做何选择做出规定。一种这样的计划或者策略是“若爱弥莉选择作贡献，则选择作贡献；若爱弥莉选择不作贡献，则选择不作贡献”。我们从逆推分析知道莱娜不会选择这一策略，但我们这里的兴趣在于对所有那些可用的策略加以描述，莱娜可根据博弈规则从中挑选出最优的策略。我们可将词语加以缩写，并用 C 表示继续，用 D 表示不继续[①]，则该策略可表述为：“若爱弥莉选 C 从而博弈位于结点 b，则选 C；若爱弥莉选 D 从而博弈位于结点 c，则选 D”，或者更简单地表述为，“在 b 选 C，在 c 选 D”，甚至写为“CD”，如果上述每一行动发生的具体情况曾

① 原文如此。怀疑原著中此处有笔误，把 contribution 错写成了 continue，如推断成立，则应译为“并用 C 表示贡献，用 D 表示不贡献”。——译者注

被解释过或是显而易见的。现在很容易看出，因为莱娜在她行动的两个结点中的每一个上都有两个可供选择的行动，所以她有四个可选择的计划或策略——“在 *b* 选 C，在 *c* 选 C”，“在 *b* 选 C，在 *c* 选 D”，“在 *b* 选 D，在 *c* 选 C” 及 “在 *b* 选 D，在 *c* 选 D”，或记为 “CC”、“CD”、“DC” 和 “DD”。在这些策略中，逆推分析和在图 3—6 中的结点 *b* 和 *c* 处的箭头表示其最优策略为 “DC”。

对于塔尼娅来说，情况要复杂一些。当轮到她行动时，根据博弈规则，行动选择的历史可以是四种可能性中的任何一种。在博弈树中的四个结点之一处，由塔尼娅行动。这四个结点的其中一个是爱弥莉选择了 C 和莱娜选择了 C 之后的那个（结点 *d*），第二个是爱弥莉选 C 和莱娜选了 D 后的那个（结点 *e*），第三个是爱弥莉选 D 和莱娜选了 C 后的那个（结点 *f*），以及爱弥莉和莱娜都选了 D 之后的那个（结点 *g*）。塔尼娅的每一种策略或完整的行动计划，都一定要对这四种场合中每一场合里她在两种行动中选择哪一个做出规定，或者说是在其四种可能的行动结中的每一个上选择其可能的两种行动之一。在有四个需要规定行动的结点以及在每个结点上有两个可选择行动的情况下，存在着 2 乘 2 乘 2 乘 2 或 16 种可能的行动组合。故塔尼娅有 16 种可选择的策略。其中一种可写为 “在 *d* 选择 C，在 *e* 选择 D，在 *f* 选择 D，在 *g* 选择 C” 或者简写为 “CDDC”。其中，我们已经将四种场合（爱弥莉与莱娜行动的历史）的顺序固定为 *d*、*e*、*f* 和 *g*。于是，利用同样的简略写法，塔尼娅可选用的全部 16 种策略为：

CCCC，CCCD，CCDC，CCDD，CDCC，CDCD，CDDC，CDDD，
DCCC，DCCD，DCDC，DCDD，DDCC，DDCD，DDDC，DDDD

在这些策略中，图 3—6 中的逆推分析和在结点 *d*、*e*、*f* 和 *g* 处的箭头表明塔尼娅的最优策略为 DCCD。

现在，我们可通过阐述每一个参与人的策略选择来表达出我们进行逆推分析的发现了——爱弥莉从其可选择的 2 种策略中选择 D，莱娜从其可选择的 4 种策略中选 DC，以及塔尼娅从其 16 种可选择的策略中选择 DCCD。当每一个参与人都沿博弈树向前看，从而预见其当前选择的最终结果时，她就可算计出其他参与人的最优策略。这种策略组合，即爱弥莉选 D，莱娜选 DC，以及塔尼娅选 DC-CD，就构成了该博弈的逆推均衡。

我们可将所有参与人的最优策略集中起来，从而找出逆推均衡中实际的博弈进行路径。爱弥莉将在开始就选 D。莱娜遵循其策略 DC 的规定选择行动 C 来应对爱弥莉的行动 D（记住，莱娜的 DC 意味着 “如果爱弥莉选了 C 则选 D；如果爱弥莉选了 D 则选 C”）。按照我们采用的规则，塔尼娅在爱弥莉选择 D 和莱娜选择 C 之后在结点 *f* 的实际行动就是表达其策略的四个字母中的第三个字母。由于塔尼娅的最优策略为 DCCD，其沿着进行路径的行动是 C。因此，实际进行路径由爱弥莉选 D，随后由莱娜和塔尼娅选 C 组成。

为了总结上述分析，我们有三种不同概念：

（1）每一个参与人可选择策略的列表。这种列表，特别是对于后面行动的参与人来说，可以是相当长的，因为必须规定他们在对应于其他参与人所有可预料得到的先前行动之场合里的行动。

（2）每个参与人的最优策略或完整的行动计划。这种策略必须规定参与人在博弈规则指定其进行行动的每个结点上的最优选择，即使在实际进行路径上许多这类结点实际上是不会达到的。这种规定事实上就是行动顺序靠前的参与人对其采取不同行动的结果的预测，因而也是他们在前面结点上算计其最优行动的重要环节。所有参与人的最优策略聚在一起就构成了逆推均衡。

（3）逆推均衡中的实际进行路径。将所有参与人的最优策略放在一起即可找出。

3.4 行动顺序的优势

在街心花园博弈的逆推均衡中，爱弥莉获得其最佳结果（支付 4），因为她可以利用第一个行动的机会。当她选择不作贡献时，她就将选择权交给了其他两个参与人——每人可以获得其次优结果当且仅当她们都选择作贡献。大多数未经深思熟虑的博弈论研究者会先入为主地认为所有博弈中都应存在**先动优势**（first-mover advantage）。然而，事实并非如此。不难想象存在这样一种博弈，其中第二个行动的人具有优势。考虑两家销售相似商品的企业之间的策略互动。如果一家企业首先公布其定价，则第二家企业会在公布自己的定价之前获知这一价格，那么，它就可以定一个稍低的价格而获得巨大的竞争优势。

先动优势来自于将其自身置于一个优势地位以及迫使其他参与人接受它的承诺能力；**后动优势**（Second-mover advantage）源于自己可对他人选择做出反应的灵活性。在一个特定博弈中，是承诺还是灵活性更为重要依赖于其特定的策略组合和支付，不存在一般性的法则。在本书中，我们将遇到两种类型优势的许多例子。一般认为并不必然存在先动优势，这与普通观念相左，但它是如此重要以至于我们感到有必要特别加以强调。

当博弈有先动优势或后动优势时，每个参与人都会试图去操纵博弈进行的顺序，以保证他自己可取得优势地位。这种操纵的策略就是策略性行动（strategic move），我们将在第 10 章中加以考察。

3.5 添加更多的行动

我们从第 3 节了解到，随着参与人人数的增加，对序贯行动博弈的分析会变得更为复杂。本节将讨论因在博弈中增加更多行动而引起的另一类复杂问题。在这里，我们只讨论仅有两个参与人且允许参与人不止一次交替采取行动的博弈。如此一来，博弈树会随之扩展开来，但博弈树后面的行动仍然由先行动者的选择决定。

许多日常游戏，诸如井字棋、跳棋、国际象棋等，都属于两人有策略地依次

轮流行动的博弈。博弈树和逆推法有助于“解出”这类博弈——确定逆推均衡结果以及参与人的均衡策略。不幸的是，随着博弈和策略复杂度的升级，寻找最优解也越来越困难。在这种情况下，人工办法已不再可行，而诸如第 2 章曾提到过的 Gambit 计算机程序就可派上用场了。

3.5.1 井字棋

我们从上面谈到的三个日常游戏中最简单的井字棋着手讨论。这里我们只考虑比平时简单的井字棋，也就是游戏参与人双方（X 和 O）都试图率先将自己的棋子连成一行、一列或一条斜线而取得胜利；他们所用的不是九格的“井”字，而是只有四格的“田”字。开始，先行者有四个可能摆放自己棋子 X 的位置，于是后行者相对于四个结点各有三种行动选择。轮到先行者走第二步时，他共有 12(4×3)个结点，对于每个结点有两个可能的行动选择。如图 3—7 所示，即使这样简单的井字棋游戏，其博弈树也相当复杂。而实际上该博弈树的规模也不算太大，因为当先行者在走完第二步后，游戏必定会结束，但是也仍有 24 种可能的终点结。

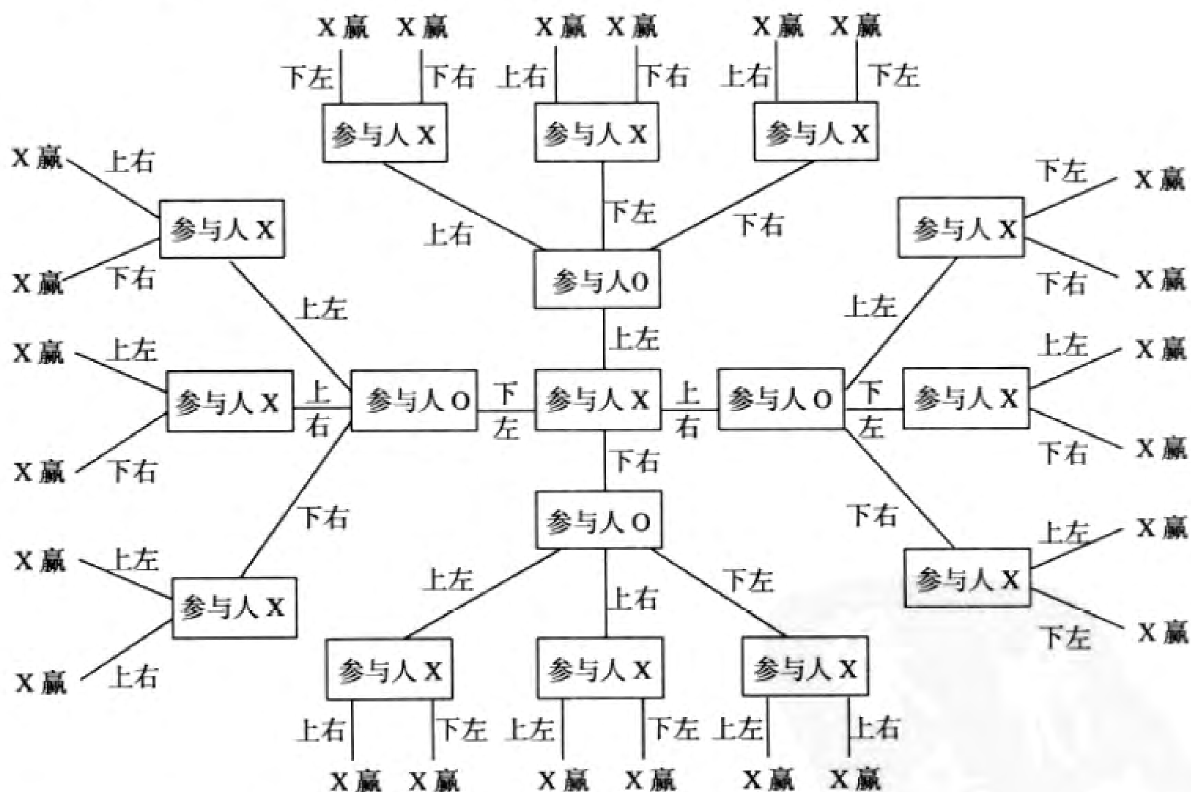


图 3—7 “田”字井字棋的复杂博弈树

举出此例仅仅是想说明即使是如此简单的（或是被简化的）博弈，其博弈树也是非常复杂的。对该博弈进行逆推分析，我们能很快得到一种均衡。逆推分析表明先行者无论怎样走第二步，其结果都一样。这里没有最优选择，每一种选择

都一样好。因此，当后行者在走第一步时，他也会发现自己不论怎样走结果都相同，也就是说他对于四个结点做出的三种行动选择毫无差别。同理，先行者在走第一步时每种选择也都一样好，因为他肯定能赢。

虽然该井字棋游戏的博弈树非常有趣，但其解却没那么有意思。先行者必定赢，所以双方做出的任何选择都不能影响最终结果。我们大多数人对九格的井字棋比较熟悉。用博弈树来分析该游戏，我们知道先行者在初始结有 9 种可能的行动选择，后行者在他的 9 个决策结各有 8 种可能的行动选择。先行者走第二步时，在 8×9=72 个结点各有 7 种可能的行动选择。于是后行者在走第二步时，在他的7×8×9=504 个结点各有 6 种可能的行动选择。这种模式会一直持续下去，直到博弈的行动组合恰好让某参与人获胜为止，但是第五步以前双方皆没有胜算。要画出该游戏的博弈树，需要一张很大的纸，或是使用很小的字体才行。

然而，大多数人都知道如何在井字棋游戏中至少打个平局。因此通过逆推分析我们能找出该游戏的一个简单均衡解，一位有经验且具策略思维的人能在很大程度上降低求解的难度。正如“田”字游戏一样，井字棋游戏博弈树中的许多路径具有相同的策略意义。9 个初始结可以分为三种类型：角落位置（4 种可能性）、边上位置（也有 4 种可能性），以及中心位置（1 种可能性）。运用这种简化博弈树的方法有助于降低问题的难度，也有助于描述最优逆推均衡策略。要明确的是，后行者只要走出正确的第一步棋并且不断阻挡先行者三子成线[3]，他就能确保至少和对手平局。

3.5.2 国际象棋

对于更复杂的游戏，诸如跳棋，特别是国际象棋，我们要找出一个完整的均衡解就更难了。有些跳棋手称，后行者总能确保平局，但并未得到证实。而国际象棋研究专家却很少有类似的说法，尽管国际象棋是一项很典型的序贯行动博弈，且适用于逆推分析法。

在国际象棋中，白棋先行，然后黑白棋双方依次轮流行动。弈棋双方都能清楚地看到对方的行动，双方均无法投机，不像在纸牌游戏中存在着洗牌或者发牌的运气问题。因此，在国际象棋游戏中，最“纯粹”的策略推理，同先前讨论的方法一样，就是向前预测你每一步会产生的结果。例如我拱兵，对手可能跳马来威胁我的车，因此，我应该在拱兵前，先用象看好对方的马。此外，国际象棋每局必须在有限的回合内结束。国际象棋规定，如果某种局面重复出现 3 次，则该局判为平局。既然要把 32 个棋子（一旦被吃掉就少于 32 个了）下在 64 个方格里的方法是有限的，那么一场比赛也就不会在不犯规的情况下无限地进行下去了。

因此，原则上国际象棋是可以用逆推法进行分析的。我们知道当双方都采用了最优策略的棋局结果：要么白棋赢，要么黑棋赢，或是平局。那么，为什么没人做这样的分析？又为什么人们继续用不同的策略玩下去，结果又场场不同呢？

答案在于，国际象棋尽管规则简单却是一种复杂且令人费尽脑力的游戏。白

棋有 20 种开局方法[4]，接着黑棋有 20 种对应的开局方法。因此博弈树初始结就有 20 条分支，接下来 20 个结点又各自延伸出 20 条分支，这样一来就已经有 400 条分支了，而每条分支连接的结点又延伸出更多的分支。一会儿的工夫，每回合可选择的下法实际上就远远超过了 20 种。通常一盘棋双方各自会走 40 步以上，估计总的可行走法约 10^{120} 种。一台运算速度为个人电脑1 000倍的超级计算机（每秒进行 10 亿次计算），需要连续工作约 3×10^{103} 年才能检查完每一步。[5]这一时间，天文学家告诉我们，比太阳蜕变成红巨星将地球吞噬进去所需的时间还要长得多。最近，有位物理学家做了一项思维实验：把整个宇宙当做一台超大型的计算机，如果把宇宙中所有的能量和物质都用来供计算机进行计算，那么自宇宙大爆炸[6]起到现在能进行多少次计算呢？答案是 10^{120} 次。很凑巧的是，倘若把自宇宙诞生以来所有的能量和物质都用来逆推分析国际象棋，我们现在仅仅处在快要完成这项工作的阶段。

很明显，要想画出完整的国际象棋博弈树，是永远不可能的。就算是描绘以某个特定下法开局的较简单的一盘棋，仅仅到第五步或第六步，博弈树就已经变得非常复杂了。我们说预测五六步是可能的，仅仅因为只要具备一点点国际象棋经验就能使游戏者认识到哪些下法明显是不明智的，并据此把它们从备选项中剔除掉。

总的来说，虽然原则上能通过逆推分析求得一个博弈的完整解，但由于整个博弈树过于复杂而使得这一方法在实践中不具可行性。面对这种情况，参与人该做什么呢？我们可以通过回顾为计算机设计国际象棋程序的历史来学习一些经验。

当计算机开始在科学和经济领域展现其强大威力时，许多数学家和计算机学家们就认为计算机将很快打败国际象棋世界冠军。然而，尽管计算机技术日新月异，进展的速度比人脑快成百上千倍，电脑战胜人脑这一天还是让我们等了很久。终于，1992 年 12 月，一台名叫德国 2 号的国际象棋计算机在一些快棋赛中击败了世界冠军卡斯帕罗夫。不过在正规赛中，即 2.5 小时内下完 40 步棋，人脑的优势维持了更久。由 IBM 赞助的一个团队投入大量精力和资源开发出了一台专业国际象棋计算机及其相应配套软件。1996 年 2 月，名为深蓝的计算机与卡斯帕罗夫展开了一场较量（六盘四胜制）。深蓝因在第一盘获胜而名声大振，但随后卡斯帕罗夫很快摸清了深蓝的弱点，改进了对攻策略，从而轻松地赢得了余下 5 盘的胜利。接下来的 15 个月中，IBM 工作组改进了深蓝的硬软件，开发出了深蓝二代，它于 1997 年 5 月的第二次六盘四胜制大赛中击败了卡斯帕罗夫。

总而言之，计算机进步的速度时快时慢，而人脑虽然在某些方面具有优势，进步的速度却不足以跟上电脑。进一步的研究表明，人脑与计算机在分析国际象棋下法的思考方法上截然不同。

如果你不能预测整个游戏的结局，那么不妨预测一下游戏的局部，例如开局后的 5 步或 10 步，然后逆推回去。游戏不会在这几步之内结束，也就是说，5 步或 10 步后还未到终点结。只有终点结才被赋予了由博弈规则所刻画的支付——从这种终点结你能逆推回去。因此，你必须赋予你所能预见到的非终点结以某种支付，但你并不能确切地知道这种支付应该是什么。由于你不能够从博弈的真正

终点结进行准确无误的逆推分析，你需要运用某种间接的方法去赋予非终点结以恰当的支付。赋予非终点结支付的规则被称为**中间评估函数**（intermediate valuation function）。

在国际象棋里，人脑与计算机都通过把中间评估函数与局部预测相结合的方法来分析问题。典型的方法如下。我们把每一个棋子都赋予一个数值，比如说兵是1，象是4，后是9。不同的棋局和棋子的组合有不同的数值。例如，两个相互协作配合行动的车比两个彼此割裂分头行动的车数值要高。我们把某步棋局上所有棋子的数值和彼此间的组合的数值加总，即为该步棋局的中间值（intermediate value）。我们并不是根据准确的逆推计算来量化和平衡某步棋局的优劣，而是根据国际象棋界的共同经验得出相关评估。于是，要想达到预期的数值，每一步都得在经过准确计算预测未来的5至6步后再做出选择。

关于开局（也就是开局的头12步左右）的中间值研究进行得最为透彻。每种开局都会产生许多不同的棋局和棋步，但是经验有助于棋手总结出对自己有利或不利的开局方法。这些经验都已被归纳进了许多关于开局的书里，所有国际象棋高手和计算机程序都能将它们牢牢地记住，并运用到实战中。

在残局期，由于棋面上棋子很少，逆推法就显得十分简单，而且棋局完整，可以推导出最终结果。例如，你可以轻而易举地推论出残局的结果：倘若一方还剩下王和车而对手仅有一个王，那么前者会赢；倘若一方剩下王和象而另一方只剩一个王，那么后者能逼平。但是，倘若再往前推一点，考虑最后的5至6步，情况就不是那么简单了，可能变得更加复杂，这时，棋手则不得不依赖间接的方法评估走法。

对中局的评估最为困难，因为此时棋局已经进展到无法在几步内简化的复杂程度，往前预测5步甚至10步都不能让你确定该从哪里着手逆推。即使你花很大的工夫计算，让自己的预测能力可以从预测5步提高到预测6步，但事实上，这些努力也不一定有助于提高你的棋艺。如果你刚巧预测到第6步的棋局对你不利，你肯定会很好地处理，但是第6步后的一些更复杂的情况会让你无法评估。换句话说，要在中局找出比较好的下法，运用中间评估函数通常比向前预测更为有效。

这就是国际象棋艺术的魅力所在。顶尖棋手具备一种直觉和本能，这种直觉和本能能够帮助他们敏锐地察觉到契机的来临，并避免陷入对方布下的巧妙陷阱，这是计算机程序难以匹敌的。计算机专家发现，要把人类本能获得和使用的模式认知（pattern recognition）技能教给计算机相当困难，譬如说识别面孔并记住其名字等。国际象棋中局下法的艺术也就跟模式认知与评估一样仍然充满神秘。这也正是卡斯帕罗夫相对德国2号和深蓝的优势所在。同时，这也解释了为什么计算机下棋程序能够在快棋战或限时比赛中打败人类，因为人脑没有足够的时间来发挥这一优势。

换句话说，顶级棋手具备一种基于经验或模式认知能力的微妙的国际象棋知识，这种知识增强了他们的中间评估能力。而计算机在处理原始数据或者进行高强度计算时与人脑相比，具有优势。有一阵子，科学家曾试图通过模拟人类思维来设计计算机程序，但这只是人工智能学界采用的方法，且并不成功。国际象棋

程序设计师渐渐意识到“把电脑就当计算机”，利用电脑强大的计算能力在更短的时间内分析计算出更多的下法将会是更明智的做法。如今尽管棋手和电脑都运用预测与中间评估函数相结合的办法，但两者结合比例不一样：棋手擅长基于国际象棋知识的中间评估，但预测能力只限于很少的几步；而计算机在复杂的评估能力上稍有欠缺，但是由于具备强大的运算能力所以能预测更远。

近年来，国际象棋计算机已经开始获取更多的国际象棋知识。1996 年至 1997 年期间，IBM 曾在国际象棋专家们的帮助下通过增强软件的中间评估能力来改进深蓝。这些专家顾问们与电脑反复对弈，寻找它的弱点，然后建议该如何修改中间评估系统以克服这些缺陷。有一段有趣的小插曲可以证实这些建议是很有建设性的。[7]国际象棋大师班杰明说：“我注意到每当深蓝通过交换兵为车打开前进道路时，从策略而言，它总是早于最优交换时机。深蓝必须明白车已经处于一个非常有利的位置，而不必立刻进行交换。它应该在打开车的前进道路之前判断车的评估点。”换句话说，深蓝的中间评估系统知道处于一个非常有利位置且行径畅通的车是有利的，因此它会毫不犹豫地立即扫清道路创造有利位置。但深蓝却没法意识到即使车的位置处理也具有“选择价值”(option value)；一旦车已经处于那个有利位置，它可随时执行换兵策略。但是如果此时兵有另外其他目的，则应该延缓换兵策略。上述这种微妙的弈棋思维是通过反复的实战经验和对棋盘上各种错综复杂的关系的认知而得来的；它对于人类棋手来说很容易获取，甚至他们常常还不清楚这一点，而对于电脑来说，把这种微妙的弈棋思维编成电脑程序比较困难。

倘若人类能够逐渐地有系统地整理出他们微妙的国际象棋知识并传授于计算机，那么人类在无法从计算机那里获得回赠的情况下将无法对战胜计算机抱任何希望。在 1997 年那场人机对战中，卡斯帕罗夫相当惊讶深蓝近乎人类甚至超越人类的弈棋能力。他甚至将计算机的某一步妙棋誉为“上帝之手”(the hand of God)。情况只可能更糟：计算机强大而快速的计算能力在不断前进中，与此同时，它们也在慢慢地获得构成人类优势的那些微妙知识。

那么计算机最终能发现自己的弱点并改良其评估能力吗？换言之，计算机能否像人类一样从自己的实战经验中获取弈棋技巧呢？如果可以，那么又是否能够把这种方式扩展到其他策略互动领域，例如商界？那么作为人类还有什么意义呢？

这些问题留给哲学家去解决吧！他们在讨论诸如人生宇宙、世间万物这样的大范畴问题方面已经有数千年的经验了，而且至今也没有任何答案。此刻，我们仅需要指出的是，不论计算机的能力有多强大，也是为人所用。毕竟，人类可以做计算机做不到的事情——人类可以拔掉插头或掐断计算机的电源，而计算机不能阻止人类。

国际象棋抽象理论告诉我们它是一项有限步的游戏，原则上可以通过逆推法求解。但是，实战则需要许多建立在经验、直觉和微妙的判断之上的“艺术”成分。这对于使用逆推法分析序贯博弈是坏消息吗？我看不是，确实，理论分析不能给我们提供答案，却能给我们许多帮助。逆推法必须结合强大的计算能力以及根据知识对中局进行评估的能力，向前几步的预测是其中重要的一部分。随着计算机能力的增强，不仅计算能力所发挥的作用在增大，就连逆推法理论的应用也

更广泛了。

我们就博弈复杂性这一点总结如下。对于一些简单游戏，我们可以不必画出博弈树，只通过口头推理找出逆推均衡。但是，对于诸如国际象棋这样复杂的博弈游戏来说，我们只能勾画出博弈树中的一小部分，因此我们必须使用以下两种方法：（1）以逆推法逻辑为基础的计算；（2）基于经验的中间评估。对于复杂程度介于两者之间的博弈游戏，用口头分析难度较大，我们可以勾画出完整的博弈树后进行逆推，有时我们可以求助于计算机来勾画并分析这种中等复杂程度的博弈树。

值得庆幸的是，我们在许多领域（诸如经济、政治、体育、商业和日常生活）中遇到的大多数博弈活动的策略的复杂程度远远不及国际象棋，可能包含数位参与人，每位参与人可以行动数次，也可能有更多的参与人，且每位参与人的行动非常多。不过，由于这些活动实质上是有序的，我们仍然可以勾画出不太复杂的博弈树。逆推法逻辑仍然有效，而且它总是有效，只要你掌握了这种逆推法思想，你无须画出博弈树就能运用逻辑思维找出解。除此以外，对于一些难度介于这一章我们已经分析清楚了的一些简单的例子和诸如国际象棋这样较难的例子之间的问题，我们可以借助例如 Gambit 这样的计算机软件进行分析；这无疑使将上述理论运用于分析实际博弈行为的前景一片光明。

3.6 逆推实验

现实中的序贯行动博弈参与人在进行逆推法推理时表现如何呢？我们还缺少许多系统性的证据，但有些实证研究所得的结论与理论预期相反。这些实验及其结果可为序贯行动博弈的策略性分析带来一些有趣的启发。

例如，许多实验是让从某一阶层或某一群自愿者中选出的两个受试者 A 和 B 进行单回合的谈判博弈。实验者提供 1 美元（或某个已知的金额），要求受试者根据以下规则进行分配：受试者 A 提出分配比例，比如说自己 75 美分，B 25 美分，若 B 接受此分配，则按此比例进行分配，若 B 拒绝，两人一个子儿都得不到。

在这个案例中，通过逆推法我们可以预测 B 应该接受 A 所提出的任何分配比例，哪怕很少，因为他的另一种选择更糟，那就是一无所获。A 要是能预见到这一点，他应该提议自己 99 美分，而只留给 B 1 美分。但是这种特殊的情况几乎永远不会发生。大多数的 A 都会提出更平均的分配，事实上，各得 50 美分将会是最常见的提议。此外，大多数的 B 拒绝 25％或者更少的份额而宁可一无所得，甚至有些拒绝 40％的份额。[8]

许多博弈论专家不认为这些实验结果有损理论的正确性，他们反驳道："可分配金额太少，参与人可能不会在乎这一点得失，B 损失 25 美分或 40 美分都无所谓，他们会放弃这一点点侮辱性的小奖励，并因此获得自我满足。如果金额变成 1 000 美元，25％也不再是个小数目，那么 B 可能会接受。"但是这种观点似

乎没有根据，因为金额更高时，实验结果也颇为相似。在印度尼西亚进行的实验虽然涉及的金额不算太大，但总量却相当于参与人三个月的收入，结果表明，尽管随着金额的增加 B 倾向于接受较少的份额，但 A 没有明显提出更不公平的分配比例。同样，在斯洛伐克的实验也揭示了缺乏经验的参与人的行动并没有受到盈利巨大变化的影响。[9]

很明显，这些受试者都不具备博弈论知识，也没有什么特别的计算能力。不过这种游戏非常简单，即便头脑最简单的人也能看出其中的逻辑而理性地进行推理决策，这可以通过实验后与参与人的问答得以证实。这些实验说明博弈论专家错误地假设了每一个参与人都只在乎金钱报酬，而并非说明逆推法无效。由于社会向其成员灌输了强烈的公平意识，这将导致大多数的 A 提出一人一半或相近的分配方案，而 B 会拒绝任何非常不公平的分配方案。对独裁博弈（dictator game）的观察很清楚地说明了这一点，也就是在这种由 A 全权决定分配而 B 没有任何选择的博弈里，许多 A 往往愿意分给 B 相当高的份额。[10]

某些实验可用来研究公平意识在像独裁博弈这类最后通牒博弈（ultimatum games）中具有的重要性。研究表明，受试者和实验者可获得的信息改变时，最终博弈结果也会有重大改变。特别是在实验设计发生变更以致实验者也无法确定谁制定分配策略，谁接受分配策略时，公平分配的可能就会显著下降。我们从新的神经经济学领域也得到了相似的结论。桑菲（Alan Sanfey）和他的同事进行了最后通牒博弈的实验，实验中他们通过核磁共振成像来观测受试者作决定时其大脑的活动状况。他们发现，当 B 拒绝不公平分配（少于一人一半）时，负责否定情感的大脑区域的行动被激活。因此在对不公平分配的拒绝中出于本能的愤怒和厌恶就溢于言表了。同时，他们也发现，当 B 得知提出不公平分配方案的 A 是人而非电脑时，他们会更加倾向于拒绝此分配方案。[11]

另一个结果与理论预测不符的实验如下。两个参与人分别为 A 和 B。实验者在桌上放面值为 10 美分的硬币，参与人可选择要或者放弃。A 先行动，若 A 选择拿走这 10 美分，游戏结束，A 将得到这枚硬币而 B 一无所获；若 A 选择放弃，实验者将增加 1 枚 10 美分的硬币，B 将面临选择获得这 20 美分或者放弃。游戏依此规则进行下去，硬币不断增加，直到某个限度（比如说 100 美分）——这一点两个参与人事先都清楚。

该游戏的博弈树如图 3—8 所示。根据博弈树形状，这类博弈通常被称为蜈蚣博弈（centipede game）。你甚至不必依赖博弈树就能对该游戏进行逆推分析。B 肯定会在最后一步行动中拿走 100 美分，因此 A 将提前，也就是在倒数第二步就把 90 美分取走。依此类推，最后的结果是 A 在第一步就拿走钱，游戏到此结束。

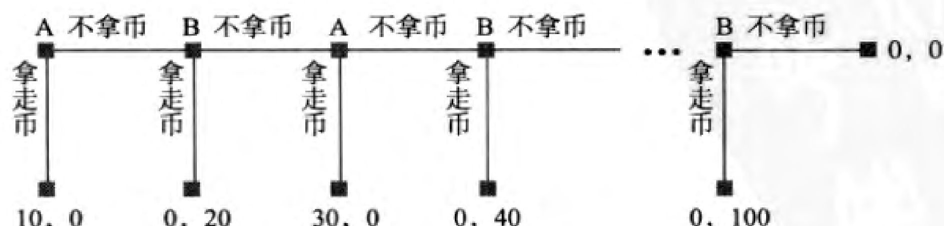


图 3—8 蜈蚣博弈

然而，在大多数实验中，游戏至少会进行几个回合。值得注意的是，通过“非理性”行动，参与人采取合谋策略要比只遵循逆推分析逻辑进行博弈获得的钱多。有时A拿走钱，有时候B拿走钱，但有时两者可能会化干戈为玉帛从而合谋解决问题。有一次迪克西特在课堂上做此实验时，该游戏就一路进行到最后一步，最后B拿走了100美分，然后他非常主动地分给A 50美分。迪克西特问A：“你们是否预先同谋？B是你的朋友吗?”A回答说：“不，我们事先互不相识，但现在我们是朋友了!”

我们发现，并不是参与人不会计算和运用博弈论逻辑，而是他们的价值体系和支付观念与博弈理论家所预测的不同。那些理论家认为，A必定在第一步就选择拿走10美分，之后游戏结束。我们在11章探讨有限次重复囚徒困境博弈时，还会遇见类似的看似与逆推分析相矛盾的合作行为。[12]

上述实例似乎说明了那些明显违背策略逻辑的行动是可以解释的，只要认识到人们并不仅仅关心自己的金钱回报，而且会在乎诸如公平等观念。但并不是所有观察到的违反逆推规则的行动都可以这样来解释。有些人的预见能力有限，也的确无法做出恰当的决策。例如，当信用卡发行者采取较低初始利率或第一年免年费的策略吸引客户时，许多人就会因此动心而忽略掉自己可能在此后付出更多。因此，逆推博弈分析在此起到的指导性或说明性的作用要比预测作用更重要。无论人们如何衡量自己的回报，掌握的逆推理论知识越多，就越容易制定好的策略，得到更高的回报。博弈论专家可以运用他们的专业知识为那些正处于复杂的策略抉择中而又无力做出最好决策的人们提供有价值的建议。

3.7 幸存者游戏策略

前面章节所介绍的例子旨在说明一些基本的概念，如决策结、枝、行动、策略，以及逆推法等。现在我们通过一个现实的例子（至少是“电视中的现实生活”）来看看它们是怎样被运用的。

2000年夏天，CBS电视台直播了“幸存者”游戏第一季，而且很快大获成功，迎来了“真实电视”时代。抛开那些复杂的细节和不相关的情节，我们首先大概了解一下游戏的规则。一群竞争者——被称为一个“部落”——被带到一个无人居住的孤岛，远离食物和庇护所。每三天他们将必须投票来决定其中一人从部落中离开（称作部落会议)，获得同伴反对票数最多的那个人就是当天的受害者。然而，在每次部落会议以前，他们会先参加一个由游戏设计者设计的考验体力或脑力的竞赛（称作挑战赛)，其胜者可以在接下来的部落会议中享有被投票淘汰出局的豁免权。当然，自己不可以投自己的否决票。最后，当只剩下两人时，七个曾被投票出局的人作为“评判团”回到游戏中，在剩下的两人中选出一个最后胜利者，他将获得100万美元。

对于所有竞争者来说，他们应采取的策略是：（1）做一个被大家认为是对“部落”在寻找食物和完成其他生存任务时都很有用的贡献者，同时避免被认为

是一个强有力的竞争者而成为被淘汰的对象；（2）结成联盟确保一定的票数来使自己不被淘汰；（3）当人数太少时，而又不得不投票淘汰某人时，背叛结成的联盟；（4）不要丧失掉好的人缘，因为被淘汰的人最后作为评判团成员有投票权。

我们来看只剩下三位竞争者时的情况，他们是鲁迪、凯利和瑞奇。在他们当中，鲁迪是最年长的一个，诚实率直，很有人缘。大家普遍认为，如果他是最后两人中的一个的话，那么他就一定会胜利获得 100 万美元。因此对凯利和瑞奇来说，在最后投票中相遇对双方都有益。但他们都不想被认为是投票淘汰掉鲁迪的人。由于只剩下三人，豁免权挑战赛中的胜者将会是投票淘汰中的绝对赢家，因为其他的两个人必须相互投反对票。于是，评判团将会知道谁该为鲁迪的离开负责。淘汰掉好人缘的鲁迪必将引起不满，而使自己在最后的投票中丧失获胜的机会。对于瑞奇来说，这的确是个棘手的问题，因为大家都知道他和鲁迪是同一联盟的。

这一轮的豁免权挑战赛是耐力的比赛：每人都得站在一个难以控制的支撑点上，倾斜着用一只手抓住中心长杆上的一个称作“安全像”的图腾。任何人的手只要离开这个图腾哪怕一会儿，就算输，谁坚持的时间最长，谁就是胜者。

比赛进行了一个半小时后，瑞奇才意识到自己的最佳策略是故意输掉这场挑战。因为，如果鲁迪赢得了这场挑战，他就会维护自己的盟友保住瑞奇——大家都知道，鲁迪是一个信守诺言的人。这种情况下，瑞奇会在最终的投票中输给鲁迪，但跟他赢得挑战保住鲁迪没有区别。如果凯利获胜，则最有可能的结果是，投票淘汰掉鲁迪，这样对她来说最有利，因为她至少还有打败瑞奇的机会，而打败鲁迪根本就不可能。这样，瑞奇赢的机会就很大了。然而，假如瑞奇自己获胜，而后投票否决掉鲁迪，他打败凯利的机会也会随他否决掉鲁迪而变小。

因此瑞奇故意输掉，而且随后在摄像头前他也清楚地阐述了自己的原因。他的判断后来被证实是正确的。凯利赢得了那场挑战并投票淘汰了鲁迪，而在最后评判团的投票中，瑞奇以一票的优势击败了凯利。

瑞奇的思考过程实质上是顺着博弈树进行逆推分析的过程。当他站在不好控制的支撑点上，倾斜着用一只手抓住“安全像”时，脑子里并没有勾勒任何的博弈树，分析完全出于本能，但他还是花了一个半小时才得出结论。了解所有关于瑞奇的思路后，我们可以清晰地画出博弈树并很快地得出答案。

博弈树见图 3—9。你可以发现，它要较前面章节出现的博弈树复杂，有更多的分支和行动，而且存在不定解，在不同情况下获胜或失败的概率需估计才能得出，而不像以前那样可以确切知晓。但是你可以看到我们是怎样得出合理假设并进行分析的。

在初始结，瑞奇要决定是继续挑战还是放弃。但无论是哪一种情况都无法预测谁是获胜者，这一点可以通过在博弈树中引入“自然”这一参与人来进行处理从而加以清楚地说明，正如我们在图 3—1 所示的抛硬币的情形下所做的一样。假如瑞奇选择继续挑战，自然会从三个竞争者中选择胜者。我们并不知道实际的可能性，但我们可以赋予某种解释一定的数值，并做出关键性的假定。假定凯利有充足的体力，而鲁迪由于年纪大而不太可能获胜。因此我们推测当瑞奇选择“继续”时，各自获胜概率为：凯利 0.5，瑞奇 0.45，而鲁迪仅为 0.05。而如果

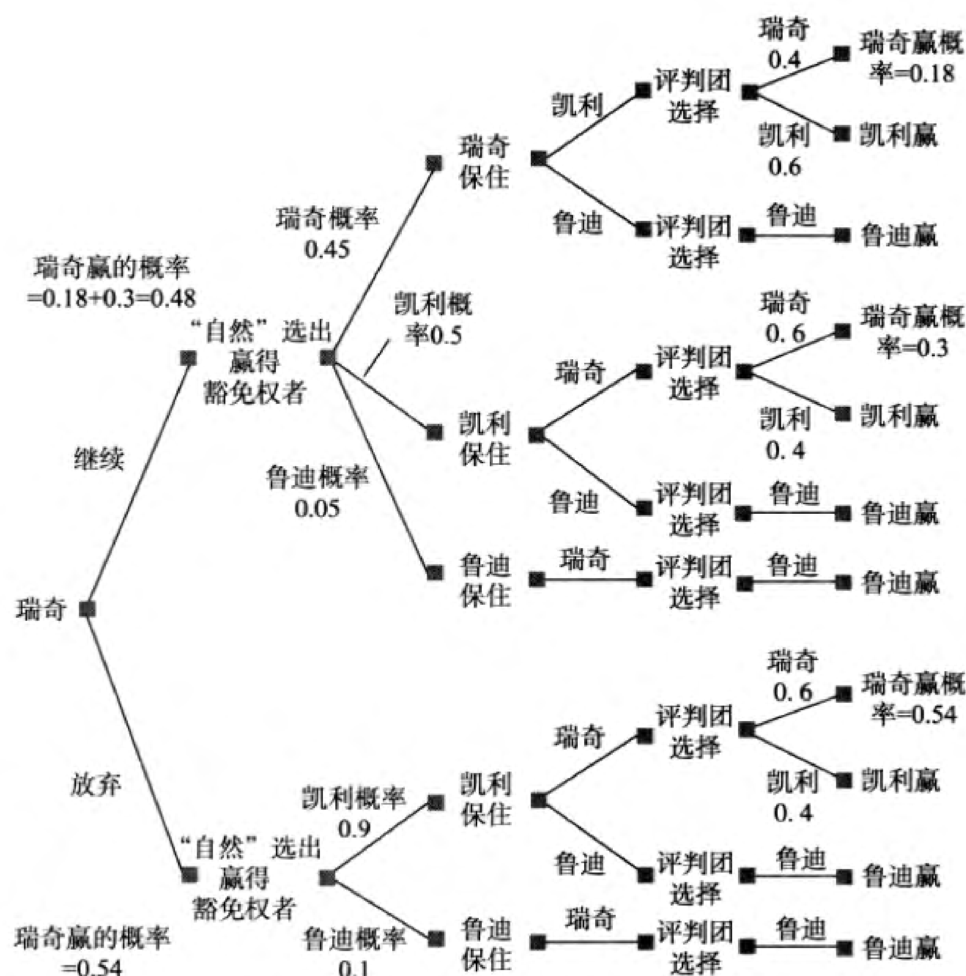


图 3—9　幸存者游戏豁免权挑战赛博弈树

瑞奇放弃挑战，自然将在剩下的两人中随机选择胜者，在这种情况下，我们假定凯利获胜的概率为 0.9，而鲁迪为 0.1。

接着博弈树随着在挑战赛中产生的三个可能胜者而延伸。假如鲁迪获胜，他会如言保住瑞奇，而评判团会投鲁迪的票使其为胜者。[13]假如瑞奇获胜，他将不得不决定是保住凯利还是保住鲁迪。如果他选择保住鲁迪，评判团则投票给鲁迪；假如他选择保住凯利，则无法确定评判团会投票给谁。我们假定，如果瑞奇选择淘汰鲁迪而得罪某些评判员，即使他比凯利更讨人喜欢，在最终评判团投票中获胜的概率也仅为 0.4。同样，假如凯利获胜，她可以选择冒着失去评判团的投票的风险而淘汰鲁迪，也可以选择淘汰瑞奇。假如她保住瑞奇，则瑞奇获得评判团投票取胜的概率会高一点，为 0.6，因为在这个情况下，他不仅获得评判团的好感，同时也不用投票淘汰鲁迪。

那么竞争者的真正支付是什么呢？毫无疑问，瑞奇和凯利都想最大化最后胜利的概率赢得 100 万美元。同样鲁迪也想获得这笔奖金，但对他来说信守对瑞奇的承诺更重要。了解各位参赛者的偏好后，现在瑞奇就可以根据博弈树进行逆推分析来决定自己最初的选择。

瑞奇知道，假如他在挑战赛中获胜（采取第一次行动后的最初路径和自然抉择），他就不得不面临只有 0.4 的最终胜率而保住凯利，而要是保住鲁迪，则意

味着毫无胜利可言。瑞奇同样也能计算出，如果是凯利在挑战赛中获胜（这种情况在博弈树的上下两分支各发生一次），她将会因为同样的理由而留下他，而他的最终胜率是 0.6。

那么瑞奇在初始结的胜率是多大呢？假如瑞奇在初始结选择放弃，那么对他来说就只有一种机会可以使自己成为最后获胜者——凯利赢得挑战赛（概率为 0.9），保住瑞奇（概率为 1），并且评判团投票给瑞奇（概率为 0.6）。因为只有上述三件事都发生了瑞奇才能获胜，那么他的总获胜概率就是这三件事情发生概率的乘积，即 0.9×1×0.6=0.54。[14]假如瑞奇在初始结选择"继续"，则对他来说有两种机会可得以最终获胜。第一种：他赢了挑战赛（概率为 0.45），淘汰了鲁迪（概率为 1），并且最终得到了评判团的投票战胜了凯利（概率 0.4）。因此，总的获胜概率为 0.45×0.4=0.18。第二种：凯利赢了挑战赛（概率为 0.5），她淘汰了鲁迪（概率为 1），但瑞奇最终得到了评判团的投票战胜了她（概率为 0.6），因此总的获胜概率为 0.5×0.6=0.3。瑞奇最后的获胜概率（在他选择继续的情况下）为这两个概率之和，即 0.18+0.3=0.48。

现在，瑞奇可以就这两种能使自己获得 100 万美元奖金的方案进行比较：一种是选择放弃，0.54 的获胜概率；另一种是选择继续，0.48 的获胜概率。根据博弈树上所有的不同方案的概率来看，瑞奇选择"放弃"会有更大获胜机会。因此"放弃"是瑞奇的最优策略。虽然结果基于我们假设的概率值，但是"放弃"仍然是瑞奇的最优策略，只要（1）凯利在瑞奇退出之后就很可能获得挑战赛的胜利；（2）瑞奇在凯利投票淘汰掉鲁迪后更易得到评判团的票。[15]

举这个例子有这么几个目的。第一个也是最主要的目的是，它向我们展示了通过运用逆推分析法，再复杂的博弈树也能分析，即使它缺乏许多准确可靠的信息。我们希望这能使你在运用这个方法时多一点信心，并且教会你如何将一些较不严谨的口头阐述推演成更准确而富有逻辑的论据。你可能会反驳说，瑞奇在推理时并没有勾勒任何的博弈树。但是了解整个系统或框架能大大简化分析过程，哪怕你身处全新而陌生的环境。因此，具有系统分析能力是必要的。

第二个目的是阐明看起来似乎荒谬的"以退为进"策略。在一些两轮制的体育比赛中参赛者也会使用此策略，例如足球世界杯。第一轮是四个球队为一组的小组赛。每组的两支领先球队晋级下一轮，他们各自的对手由事先规定好的模式决定。例如 A 组排名第一的球队对 B 组排名第二的球队，依此类推。在这种情况下，对于某个球队来说最优策略可能就是在第一轮比赛中以第二的排名晋级下一轮，因为这样一来，由于某种原因，所面对的对手要比在第一轮比赛以第一的排名进入第二轮比赛时所面对的对手更容易击败。

3.8 总　　结

序贯行动博弈需要参与人在每次采取行动之前都考虑决策的后果。分析单纯的序贯行动博弈通常需要建立博弈树。博弈树由若干个结和分支组成，代表每个

参与人在每次可以行动的机会中所有可行的选择。博弈树也给出了参与人在各种可能结果下所获得的支付。参与人策略是一套完整的计划，它说明在其他参与人可能先采取的各种行动组合下，自己该如何应对。序贯行动博弈中的均衡指的是逆推均衡，在此均衡中参与人的均衡策略就是通过预测未来局势为现在做最佳决策。这个过程被称作逆推法，或是逆推归纳法。

不同类型的博弈对不同参与人来说有不同的优势，例如先动优势。参与人或行动的增加只是扩展了序贯行动博弈的博弈树，但并没有改变推理的过程。在有些情况下，我们没有足够的空间和时间来完整地勾勒出某个博弈的博弈树。对此类博弈，我们通常是通过简化博弈树（识别出行动在策略上的相似性）或通过简单的逻辑思考。

国际象棋，一种最具策略性的序贯行动博弈，理论上可以通过逆推法进行分析。然而，由于该博弈过于复杂，实际上不可能完整地画出它的博弈树和整理出对它的分析过程。在实际的博弈中，艺术性（对模式与机会的识别）和科学性（向前预期并估计出某步棋可能的结果）对参与人决定如何行动都很重要。国际象棋选手在知识和本能的运用上仍有优势，但在使用逆推法进行计算时计算机则比人更快更准确。随着计算机功能的不断完善，运用逆推法更全面地分析象棋已慢慢成为可能。

某些序贯行动博弈实验似乎显示实际参与人存在非理性行为，又或者理论没能准确地预测未来行动。但反驳者指出对不同结果的实际偏好问题是复杂的，当获悉参与人偏好时可以运用策略理论找出最优行动。

关键术语

行动结（action node）
逆推归纳法（backward induction）
分支（branch）
决策结（decision node）
决策树（decision tree）
均衡进行路径（equilibrium path of play）
扩展式（extensive form）
先动优势（first-mover advantage）
博弈树（game tree）
初始结（initial node）
中间评估函数（intermediate valuation function）
行动（move）
结（node）
进行路径（path of play）
剪去（prune）
逆推（rollback）
逆推均衡（rollback equilibrium）
根（root）
后动优势（second-mover advantage）
终点结（terminal node）

习　题

1. 假设两个参与人 A 和 B 参加一个序贯行动博弈，A 先走，B 后走，每个参与

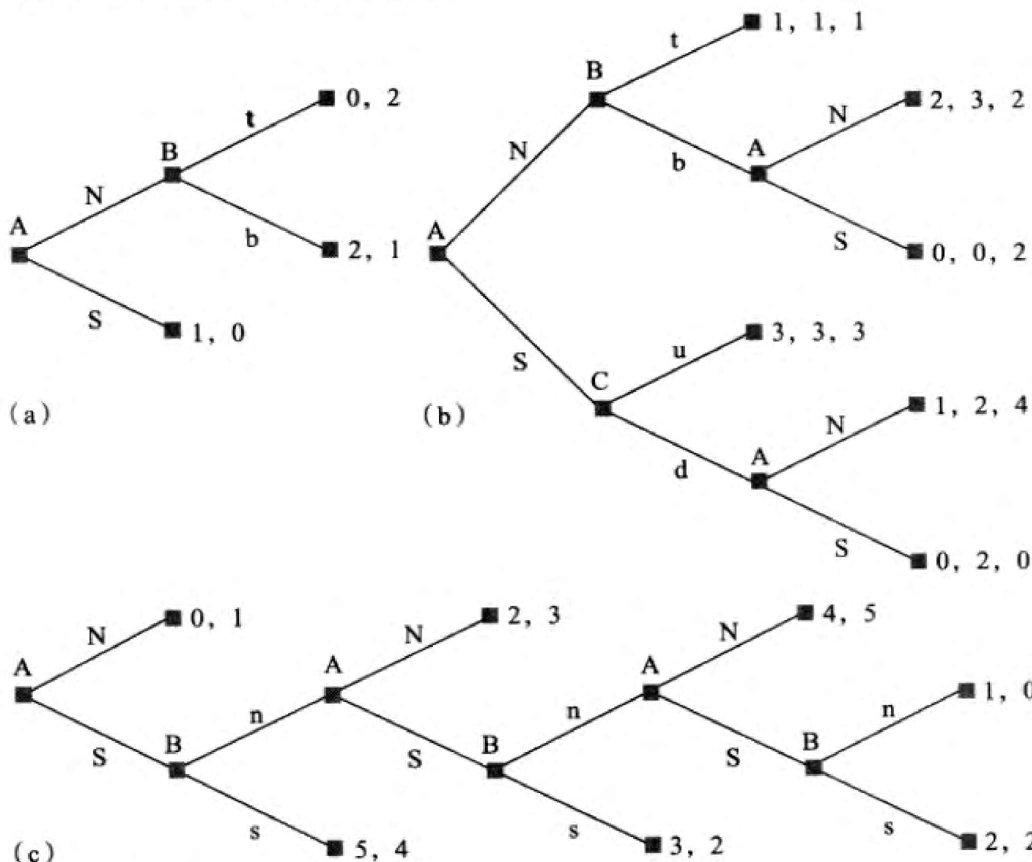
人只有一次行动机会。

(a) A在每一结点可采取上或下两种行动，B在每个结点可采取上、中或下三种行动。请画出博弈树，且问有多少决策结和终点结？

(b) A和B在每个结点可采取坐、站或跳三种行动。请画出博弈树，且问有多少决策结和终点结？

(c) A在每个结点可采取北、南、东、西四种行动，B在每个结点可采取停留或离开两种行动。请画出博弈树，且问有多少决策结和终点结？

2. 运用逆推法找出下列博弈的均衡。



请分别指出每个参与人有多少可行的纯策略并描述这些纯策略。

3. 请找出习题2中每个博弈的均衡结果以及每个参与人的完全均衡策略。

4. “在序贯行动博弈中，先手定会获胜。”这句话是否正确？请简要陈述你的原因，并举例说明。

5. 假设空中客车与波音两家公司在研发新型商业客机方面展开了激烈的竞争。波音公司在研发过程中已处于领先地位，而空中客车公司正考虑是否参与这场竞争。假如空中客车不参与竞争，那么它的收益为零，而波音公司将会获得垄断地位，获得10亿美元的收益。假如空中客车公司决定参与竞争，则波音公司就不得不决定是与空中客车公司和平竞争，还是打价格战。如果和平竞争双方各自可获得3亿美元的收益。而如果打价格战，则客机价格下滑，双方都无法收回研发成本，各损失1亿美元。请画出博弈树，找出逆推均衡并描述两家公司的均衡策略。

6. 我们不必画出完整的博弈树，便可以通过逆推法分析图3—8中的蜈蚣博弈。

在正文中讨论过的博弈中，A和B两个参与人都轮流有机会拿走或放弃桌上不断累积的10美分硬币（总数不超过10枚）。假定博弈规则发生变化：

(a) 参与人A每给对手一次机会，可获得5美分的奖励。找出每个参与人的逆推均衡策略。

(b) 两个参与人A和B有两轮机会。在第一轮，A最多只能拿走5枚硬币，而B不能超过9枚。找出每个参与人的逆推均衡策略。

(c) 两个参与人A和B有两轮机会。在第一轮，A最多只能拿走5枚硬币，而B不能超过4枚。找出每个参与人的逆推均衡策略。

7. 两个参与人艾米和贝斯，从艾米开始，轮流选数。每个参与人可选择1到10之间（包括1和10）的任何数。两个参与人所选数字相加达到100时，博弈结束。假设有这样两种结果：(1) 所选数字让总额正好为100的参与人为胜者；(2) 所选数字让总额等于100或大于100的选手为负者。对这两种情况，分别回答以下问题：

(a) 谁会获胜？

(b) 每个参与人的最优决策是什么？

8. 假设巨人、太阳神、弗里达三大百货公司正考虑在波士顿两个新的大型购物中心中的一个开设分店。其中，城市购物中心靠近人口密集的富人区，规模不大，最多只能以两家大百货商场为龙头。而郊区购物中心地处较远的郊外，相对较穷，能以三家百货商场为龙头。三家百货公司都不想在两个地方同时开店，因为顾客有相当部分重复，两处都开店无疑是同自己竞争。每家百货公司都不愿意在一个地方独家经营，因为拥有多家商场的购物中心能够吸引更多的顾客，顾客总量的增加自然会使商场利润增加。此外，它们都偏向争夺富人群体的城市购物中心。所以它们必须在城市购物中心（如果这个尝试失败了，它们将会尝试在郊区建立商场）和郊区购物中心（不争取城市市场，而直接进入郊区市场）之间做出选择。

在这个案例中，百货公司将5种可能结果按等级排列如下：5（最好），和另一家公司在城市购物中心；4，和一家或两家公司在郊区购物中心；3，在城市购物中心独家经营；2，在郊区购物中心独家经营；1（最坏），在尝试进入城市市场失败后在郊区独家经营，而此时其他非百货业商店已签约获得郊区购物中心的最好地盘。

三家百货公司因管理结构各不相同，所以做新购物中心扩展市场文书工作的快慢也不相同。弗里达公司动作最快，其次是巨人公司，最后是太阳神公司，它在准备选址方案方面效率最低。当三家公司都提出申请后，由商场来决定哪一家公司可以进入。因为品牌效应，巨人公司和太阳神公司都有潜在的客户，因此，商场首选其中的一家或两家，然后才是弗里达公司。这样，如果三家都申请进入城市购物中心，弗里达肯定不会获得一席之位，尽管它最先行动。

(a) 画出购物中心选址博弈的博弈树。

(b) 阐述用逆推法简化博弈树的过程，并利用简化后的博弈树找出逆推均衡。通过每家公司采取的（完全）策略描述该均衡。每家公司在达到逆推均

衡时会有怎样的支付？

9. 在华盛顿，有两个备受争议的提案 A 和 B。议会赞成 A，总统赞成 B。两个提案并不是非此即彼，其中一个可能，或都可能，又或都不可能成为法律。因此出现了四种可能结果，双方的支持度如下表所示，其中较大的数值表示其更受支持。

结果	议会	总统
A 成为法律	4	1
B 成为法律	1	4
A 和 B 都成为法律	3	3
A 和 B 都不成为法律（维持现状）	2	2

(a) 博弈的过程如下。首先，议会决定是否通过一个支持 A 或 B 提案或 A 和 B 两个提案的法案。然后，总统来决定是通过还是否决此法案。而议会没有足够的票数来推翻总统的否决提案。画出博弈树，找出逆推均衡。

(b) 现在假设该博弈规则仅有一处改变，即总统获得了一种额外否决权。这样一来，当议会通过了支持 A 和 B 提案的法案时，总统不仅可以完全通过或否决该法案，而且也可以选择否决掉该法案支持的其中的一个提案。画出新的博弈树，找出逆推均衡。

(c) 解释以上两个均衡不同的原因。

10. 相比于图 3—9 的幸存者博弈，这里不同的是，你只需要使用具有一般意义的符号来表示不同概率。设瑞奇选择“继续”后赢得豁免权挑战赛的概率是 x，凯利是 y，鲁迪是 $1-x-y$；同样，设瑞奇选择“放弃”后，凯利获胜的概率是 z，鲁迪是 $1-z$。此外，设瑞奇赢得挑战赛又投票否决掉鲁迪后被评判团选择的机会是 p，而凯利赢得挑战赛并淘汰鲁迪后瑞奇获胜的机会是 q。继续假设下去，若鲁迪获得挑战赛胜利，他保住瑞奇的概率是 1，并且他在最后二选一中获胜的概率也是 1。

(a) 求当瑞奇选择“继续”时，赢得 100 万美元概率的数学表达式，并用 p，q，x 和 y 表示。求当他选择“放弃”时获胜概率的数学表达式。你能在目前的信息水平下确定瑞奇的最优策略吗？

(b) 3.7 节中的讨论表明“放弃”是瑞奇的最优选择，只要（1）当瑞奇放弃后，凯利很可能获得挑战赛的胜利；（2）瑞奇在凯利投票淘汰鲁迪以后，比凯利更易获得评判团的投票。用概率符号（p，q，x，y）表示这两种条件。

(c) 假设在（b）中的两条件成立，证明瑞奇的最佳选择是“放弃”。

【注释】

[1] 在后面的章节中，我们将改变该博弈的规则——行动的顺序和支付——并考察这种改变是如何影响结果的。

[2] 回忆在第 1 节中关于博弈树的一般讨论，对于序贯行动博弈，一般对应于参与人行动的顺序列出支付；然而，为避免模棱两可或让表述更加清楚，明确地规定好顺序也是一种好的做法。

[3] 如果先行者把第一步下在中心位置，后行者就必须把自己第一步下在角落位置。接下来，后行者要确保平局就得占据先行者试图想连接的横、竖、斜的第三格位置。如果先行者把第一步下在边上位置或角落位置，后行者就必须把自己第一步下在中心位置，接着采取上述相同的阻挡技巧，同样也能确保平局。要注意的是，如果先行者选择的是角落位置，后行者选择的是中心位置，接着先行者选择的是他自己第一步棋的对角位置，那么后行者要想确保平局就一定不能下在角落位置。

[4] 白棋可以移动 8 个兵中的任意一个向前一格或两格，或者跳马（到车 3 或象 3 的方格内）。

[5] 这种计算只需进行一次，因为解完该博弈后，任何人都能运用该解，而实际上就没有人再需要玩此游戏了。每个人都会知道白棋是否有必胜策略，黑棋是否有办法保证至少和局。棋手就会投币决定谁持黑谁持白，知晓结果后握手回家。

[6] "The Physics of Information," *The Economist*, June 6, 2002.

[7] 该说明摘自 1997 年 5 月 18 日《纽约时报》的一篇由 Bruce Weber 撰写的文章 *What Deep Blue Learned in Chess School*。

[8] 该博弈和相关博弈的详尽阐述请参阅 Richard H. Thaler, "Anomalies: The Ultimate Game," *Journal of Economic Perspectives*, Vol. 2, no. 4, fall 1988, pp. 195 - 206, 以及 Douglas and Charles A. Holt, *Experimental Economics*, Princeton University Press, 1933, pp. 263 - 269。

[9] 印度尼西亚的实验结果见 Lisa Cameron, "Raising the Stakes in the Ultimatum Game: Exprimental Evidence from Indonesia," *Economic Inquiry*, Vol. 37, no. 1, January 1999, pp. 47 - 59。Slonim 和 Roth 称他们的实验结果与 Cameron 的类似，而且还发现随着盈利的增加，拒绝分配的次数有所减少。参见 Robert Slonim and Alvin Roth, "Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic," *Econometrica*, Vol. 66, no. 3, May 1998, pp. 569 - 596。

[10] 有人认为，公平这一社会规范在全体社会成员参与的不断发展的博弈中可能的确很有价值。看重公平的参与人会降低交易成本和冲突的代价，从长远来看，这是有利于社会的。这个问题我们将在第 11 章、第 12 章里讨论。

[11] Alan Sanfey, james Rilling, Jessica Aronson, Leigh Nystrom, and Jonathan Cohen, "The Neural Basis of Economic Decision-Making in the Ultimatum Game," *Science*, Vol. 300, June 13, 2003, pp. 1755 - 1758.

[12] 再一次，有人会想如果每一步增加的金额是 1 000 美元而非 10 美分，那么情况会怎样呢?

[13] 从学术的角度分析，鲁迪在豁免权挑战赛获胜后的行动结处面临着保住瑞奇或凯利的选择。我们指出鲁迪的唯一选择必定是保住瑞奇，因为每一个人都认为他根本不可能选择保住凯利（鉴于鲁迪-瑞奇联盟)。同样，评判团在最终行动结处也面临着在鲁迪与瑞奇之间选择。此外，先前的结论是在这种情况下鲁迪会赢。

[14] 需要了解和重温有关概率合并规则的读者可以参考第 7 章的附录。

[15] 掌握了概率论知识的读者可以运用更具一般性的符号而非具体数字来求解该博弈，例如本章习题 10。

第 4 章　纯策略同时行动博弈（Ⅰ）：离散策略情形

第 2 章曾提及，如果参与人在不清楚对手所做选择的情况下采取行动，这种博弈就被称为同时行动博弈。很显然，参与人在同一时间选择他们各自的行动是一种同时行动博弈。此外，即使这些行动选择不是同时进行的，但参与人在分别选择各自的行动时，不清楚其他人所做过的和将要做的选择，那么这样的博弈也是同时行动博弈（因此，同时行动博弈常常又指不完美信息或不完美知识博弈）。本章将重点讨论参与人同时行动的博弈：考虑各种不同类型的同时行动博弈，引入纳什均衡的概念，并且研究单一均衡、多重均衡和无均衡等各种情形。

许多大家熟知的策略场合都可用同时行动博弈加以描述。例如，电视机、立体声音响或汽车的制造商们在不清楚竞争厂商的有关产品决策信息时，决定各自的产品设计和特征；在美国大选中，选民的投票是同时进行的，因为他们都不知道其他人给谁投了票；在罚点球时，守门员和对方射手就必须同时做出决策——守门员不可能等到球被踢出后再决定往哪边扑救，因为这样做就太晚了。

在同时行动博弈中，某个参与人在决策时显然不知道其他人的行动，而且他也不能预见其他人对自己的行动将要做出的反应，因为行动的选择是同时进行的，其他人也不清楚他所选的行动。然而，各个参与人都在猜测对手们的行动，同时对手们也在权衡自己的行动。尽管这种循环使得同时行动博弈的分析比序贯行动博弈更难以理解，但分析起来并不困难。在本章中，我们将介绍一个简单的均衡概念，它对这种博弈将会有很强的解释力和预测力。

4.1 离散型策略同时行动博弈的描述

在第 2 章和第 3 章中，我们曾强调策略是指一套完整的行动计划。而在纯策略同时行动博弈中，各参与人至多只有一次行动的机会（虽然这次行动可能包含许多步骤）；如果参与人有多次行动的机会，那么这些机会则是构成序列的元素。因此在同时行动博弈中，策略和行动没有明显的区别，并且此时这两个概念可以互为同义词使用。但有一点值得注意：策略可以是在行动集上的随机选择。例如，在体育竞技中，参赛者或参赛队可能有意识地随机选择行动，让对手猜疑不定。这种随机选择的策略被称为**混合策略**（mixed strategies），我们将在第 7 章和第 8 章中详细讨论。在本章中，我们只讨论最初定义的行动，即**纯策略**（pure strategies）。

在许多博弈中，每个参与人只拥有有限个离散型的纯策略供其选择，例如，在篮球比赛中的运球、过人或投篮。而在另一些博弈中，每个参与人的纯策略可以是一个连续区域中的任意一个数，例如，厂商制定的价格。[1]这种区分对理解均衡的一般概念是没有影响的，但当策略是离散型时，意思更容易表达一些；同时，求解连续型策略的博弈将需要稍微高深一点的方法。因此，在本章中，我们的讨论仅限于离散型纯策略，至于策略是连续型变量的情形，我们将在第 5 章中加以介绍。

离散型策略同时行动博弈通常可用**博弈表**（game table）［也称为**博弈矩阵**（game matrix）或**支付表**（payoff table）］来描述，这张表称为博弈的**规范式**（normal form）或**策略式**（strategic form）。不论参与博弈的人数为何数，都可以用博弈表进行描述，但是表的维度必须等于参与人的个数。对于两人参与的博弈而言，博弈表是二维的并且和平面数据表相似。表中行和列上的文字分别表示了第一个和第二个参与人可选的策略，所以表的规模是由参与人策略的数量决定的。表中的每个单元格列出了相应策略组合下各个参与人的支付。三人参与的博弈需要一个三维表来刻画，我们在本章中稍后考虑这个问题。

我们用表 4—1 所描述的简单博弈来解释支付表的概念。该博弈没有什么具体的内容，这样便于我们专注地理解概念而不为“故事”分心。参与人分别被命名为“行”和“列”。“行”有四个选择（策略或行动）：顶、高、低、底；“列”有三个选择：左、中、右。“行”和“列”的每个策略所构成的组合都会产生一个可能的博弈结果，同时每个结果相对应的支付都标注在相应行和列的单元格中。按照惯例，两个支付数值中的第一个是“行”的支付，第二个数值是“列”的支付。例如，若“行”选择“高”，同时“列”选择“右”，“行”的支付是 6，“列”的支付是 4。

表 4—1　　同时行动博弈表

		列		
		左	中	右
行	顶	3，1	2，3	10，2
	高	4，5	3，0	6，4
	低	2，2	5，4	12，3
	底	5，6	4，5	9，7

在一些博弈中，特别是在体育竞技中，博弈双方的利益是完全相对的。对于参与人的每个策略组合而言，一个参与人的支付就是另一个人支付的相反数。正如第 2 章所述，我们称这种博弈为**零和博弈**或**常和博弈**。

在零和博弈中，我们可以通过只标注一个参与人（通常是行参与人）的支付来简化博弈表，那么另一个参与人（列参与人）的支付也隐含在表中了。我们用简化后的（美式）橄榄球中的单次对局来说明这种简便的标注法，如表 4—2 所示。攻方的球队试图向前移动足球以提高其射门得分的概率，它有四种可能的策略：跑位以及短、中、长三种不同距离的传球。守方可以采用以下三种策略中的一种以牵制攻方：跑动防守、传球防守、迅速四分卫。这是一个零和博弈，攻方试图得分，而守方全力阻止他们的进攻。假设我们对攻守双方的实力有足够的信息，以至于可以计算出完成不同对局的概率以及各种策略组合下的期望得分。例如，攻方选择中传，同时守方采取传球防守，我们估算攻方的支付为 4.5 码[2]，而守方的支付为－4.5 码，但守方的支付数值没有在表中标注。类似地，其他单元格也只标出了攻方的支付，并且隐含了守方的支付（等于攻方支付的相反数）。

表 4—2　　零和的同时行动博弈表

		守方		
		跑动防守	传球防守	四分卫
攻方	跑位	2	5	13
	短传	6	5.6	10.5
	中传	6	4.5	1
	长传	10	3	－2

4.2　纳什均衡

为了分析同时行动博弈，我们有必要考虑参与人是如何选择他们的行动的。回到表 4—1 所示的博弈表，我们重点考察其中的一个结果——“行”选择“低”，“列”选择“中”；“行”的支付是 5，“列”的支付是 4。每个参与人都试图选择能获得最高支付的行动，并且在博弈的结果中，给定对手的选择，每个参与人确实都是如此决策的。给定“行”选择“低”，“列”能否通过选择不是

"中"的行动而提高支付呢？不能，因为"列"选择"左"只能得到2，选择"右"只能得到3，这两个支付都不比他选择"中"时的支付4高。因此，"中"是"列"对"行"的行动"低"的**最优反应**（best response）。反过来，当"列"选择"中"时，"行"能否通过选择不是"低"的行动而变得更好些呢？答案也是否定的。因为"行"选择"顶"的支付是2，选择"高"的支付是3，选择"底"的支付是4，这些行动相对于"低"（得到5）而言，都不能使"行"的情况变好。因此，"低"也是"行"对"列"的行动"中"的最优反应。

"行"的选择"低"、"列"的选择"中"都具有如下性质：他们的选择都是对其他人行动的最优反应。一旦他们做出这样的选择，他们就没有动机独自改变行动。由非合作博弈的定义可知，参与人都是各自独立做出决策的，所以单边改变行动是博弈双方可以思考和选择的全部。正因为没有任何一方想做出如此改变，所以很自然，这种状态就可称为一种均衡，确切地说，是**纳什均衡**（Nash equilibrium）。

更加规范的定义可表述如下：纳什均衡[3]指的是博弈中的一个策略组合（每个参与人选择一个相应的策略），在其他参与人都坚守这个策略组合中的策略不变的情况下，没有参与人可以通过改变自己的策略而得到一个更高的支付。

4.2.1 纳什均衡概念的进一步解释

为了更好地理解纳什均衡的概念，我们再回到表4—1所示的博弈表。考察其中一个不是（低，中）的单元格——"行"选择"高"且"列"选择"左"，这样的策略组合是纳什均衡吗？不是，因为如果"列"选择"左"，"行"选择"底"可以得到5，而选择"高"只能得到4，所以"行"通过这样的调整而变得更好。与此类似，（底，左）也不是纳什均衡，因为"列"可以通过改选"右"而将其支付从6提高到7。

纳什均衡的定义不要求构成均衡的策略严格优于其他可选的策略。表4—3与表4—1基本一致，但其中"行"在（底，中）上的支付变成了5，与他在（低，中）上的支付相同。当"行"选择"低"且"列"选择"中"时，"行"改变行动不能使其情况变得更好。所以当博弈结果是（低，中）的时候，博弈双方都没有动机改变行动，这样便说明（低，中）是一个纳什均衡。[4]

表4—3　　　表4—1中博弈的变型

		列		
		左	中	右
行	顶	3，1	2，3	10，2
	高	4，5	3，0	6，4
	低	2，2	5，4	12，3
	底	5，6	5，5	9，7

值得强调的是，纳什均衡并不是指所有参与人的支付最优。如表 4—1 中，策略组合（底，右）的支付为（9，7），这比纳什均衡时的支付（5，4）要好。但是，当参与人独立决策时，策略组合（底，右）是不可维持的。给定“列”选择“右”，“行”若偏离“底”并选择“低”，他可以得到 12 而不是 9。只有防止“背叛”的合作行为，才能获得共同最优的支付（9，7），我们随后将在本章中讨论这种类型的行为，并且将在第 11 章中做更进一步的分析。在这里，我们仅仅指出纳什均衡可能不是所有参与人的利益最优的结果。

为加深对纳什均衡的理解，我们再回到表 4—2 所示的橄榄球比赛博弈。若守方选择传球防守，那么攻方的最优选择就是短传（这样可以得到 5.6 码而不是 5 码、4.5 码或 3 码）。反过来，如果攻方选择的是短传，那么守方的最优选择是传球防守——这可以使攻方只前进 5.6 码，而守方在选择跑动防守和四分卫时，将分别输掉 6 码和 10.5 码（记住，在零和博弈中，每个单元格中的数值只是行参与人的支付。因此，列参与人的最优选择是在单元格中数值最小而不是最大的时候取得）。在这个博弈中，策略组合（短传，传球防守）是一个纳什均衡，攻方所得到的支付是 5.6 码。

如何发现博弈中的纳什均衡呢？你可以逐个验证各个单元格所对应的策略组合是否符合纳什均衡的定义。这种**逐个单元格检查**（cell-by-cell-inspection）或枚举（enumeration）的办法虽然简单，但非常繁冗，除非博弈本身很简单或有计算机程序辅助，否则这种方法是不可行的。值得庆幸的是，还有许多适用于特定博弈类型的其他方法，不仅可以更加快速地找到纳什均衡，还可以使我们对信念和选择的形成过程有更好的理解。我们将在随后的小节中介绍这些方法。

4.2.2 纳什均衡：一个信念和选择的体系

在进一步讨论纳什均衡概念的应用之前，我们在这里要澄清几个容易让人误解的问题。我们曾提到，在纳什均衡中每个参与人的选择都是对其他人策略的最优反应，但是参与人都是同时做出这些选择的，那么他们怎么能够对没有发生的事情（至少是在不知道其他参与人的选择的情况下）做出反应呢？

人们无时无刻不在进行同时行动博弈，并且需要经常做出决策。因此，人们必须要找到一种替代真实知识或直接观察别人行动的方法。参与人可以盲目地进行猜测，并且期望这些猜测是有效的。然而，他们还有更加系统的方法来分析或计算其他人的策略选择。其中一种方法就是依据经验和观察——如果某人经常与相同或相似的其他人一起进行相同或相似的博弈，那么他就会对其他人的行为有很好的了解，那些非最优的选择将是不会持久的。另一种方法则是使用“想其他人之所想”的逻辑，你可以设想自己是其他参与人并且思考他们是怎么想的，当然，其他参与人也会站在你的立场上考虑你是怎么想的。这种逻辑看起来像个循环，但是有若干种方法可以解除这种循环，我们将在随后的小节中通过例子来说明。纳什均衡可以理解为这种“想其他人之所想”过程的顶点，其中每个参与人都正确地分析出其他人的选择。

无论是通过观察、逻辑推演还是其他什么方法，博弈的参与人都获得了关于其他人在同时行动博弈中选择的认识。要找到一个词语来概括这个过程或这个过程的结果是不容易的，它既不是期望，也不是预测，因为其他参与人的行动不是在未来出现的，而是与你的行动同时出现的。博弈论学者经常使用“信念”（belief）一词来刻画这个过程或结果，但这个词并不完美，因为相比于实际想说的意思，它包含了过多的确定成分。事实上，我们将在第5章中允许信念保持一定的不确定性。在没有更好词语表达的情况下，信念一词还是很有说服力的。

下面我们从信念的角度来理解纳什均衡。我们曾定义纳什均衡为这样一个策略组合：其中每个参与人的策略都是对其他人策略的最优反应。如果某个参与人不知道其他参与人的真实选择，但具有关于他们选择的信念，那么在纳什均衡中，这些信念必须是准确的，即你所认为的和别人所做的是一致的。因此，我们可以将纳什均衡用另一种等价的方式来定义：有这样一个策略组合，每个参与人都选择了一个相应的策略，并且具备如下性质：(1) 每个参与人都对其他参与人的策略有正确的信念；(2) 给定每个参与人关于其他参与人策略的信念，自己所选择的策略是最优的。

这样理解纳什均衡有两个优点：第一，最优反应的概念不再有逻辑上的缺陷。每个参与人都选择了基于自己业已形成的关于其他参与人行动的信念的最优反应，而不是对没有观察到的其他参与人的真实行动的最优反应。第二，在第7章和第8章关于混合策略的讨论中，参与人A在其策略集上的随机选择可以解释为其他参与人对A策略选择的信念的不确定性。现在，我们还是继续同时使用这两种对纳什均衡的解释。

你可能会认为正确信念的形成和最优反应的计算对人们来说很困难，我们将在后面的章节中讨论此类看法以及关于纳什均衡的经验和实验证据，其中第5章将讨论纯策略的情形，第7章、第8章将讨论混合策略的情形。而在本章中，正如布丁的味道如何要从吃中来体会一样，我们将在应用中说明和介绍纳什均衡的概念，我们希望在应用中学习能比抽象的讨论更有助于你掌握纳什均衡的优点和弊端。

4.3 占优策略选择

在一些博弈中，某个策略总是优于或劣于另一个策略。如果博弈具有这样的特殊性质，就可用一种专门的方法来寻找纳什均衡，并且对纳什均衡的解释也变得简单了。

大家熟知的**囚徒困境**（prisoners' dilemma）博弈就是一个很好的例子，这类故事也经常出现在电视节目《纽约重案组》（*NYPD Blue*）中。假设一对夫妇被怀疑共同谋杀了一名年轻女子，他们被带到警察局，侦探西波维茨（Sipowicz）和克拉克（Clark）将这两名嫌疑犯分别关在不同的审讯室里，并且进行隔离审问。警方暂无确凿的证据证明是这对夫妇所为，但有一些迹象表明他们曾绑架过

受害人。侦探向每个嫌疑犯都讲明这样的政策：即使他们两个谁都不认罪，他们也将被指控犯有绑架罪，分别面临 3 年徒刑；此外，侦探们分别告诉丈夫和妻子，他们已经“知道”整件事情的原委，并且也“知道”有一人是被另一人胁迫参与犯罪的。他们还被暗示，如果一个人认罪，另一个人抵赖，一旦记录下来，那个唯一的认罪者的刑期将大大缩减（几乎在每一集中，此刻荧幕上出现的都是摆放在书桌上的黄色法律便笺和一支铅笔）。最后，这对夫妇还被告知如果他们两人都认罪，刑期将可协商，且比他们中只有一人认罪而另一人抵赖时的抵赖者的刑期要短。

在这样一个由这对夫妇参与的同时行动博弈中，他们可以选择认罪或抵赖。他们都明白，如果没人认罪，他们分别会因绑架罪而服刑 3 年。同时，他们也清楚，若有一人认罪，认罪的一方将因与警方的合作而得到一个较短的 1 年刑期，而抵赖的一方将至少被判 25 年。如果双方都选择认罪，他们知道，可以通过谈判，每人各被判 10 年。

将这个博弈的策略选择和结果也用博弈表概括，如表 4—4 所示。为了更好地表达两个参与人在相互对局中的关系，“认罪”与“抵赖”又被称为“背叛”与“合作”：“背叛”是指背离双方默许的安排，“合作”是指采取帮助对方的行动（而不与警察合作）。

表 4—4　　　　囚徒困境

		妻子	
		认罪（背叛）	抵赖（合作）
丈夫	认罪（背叛）	10，10	1，25
	抵赖（合作）	25，1	3，3

这里的支付就是在各种情形下该夫妇各自被判入狱的年数，所以更小的数对于参与人来说才是更好的。在这一点上，这个例子同许多“支付越大越好”的博弈是不同的。借此机会，我们也提醒读者，“支付越大越好”并不总是正确的。在支付数值表示参与人结果的排序时，人们通常使用 1 表示最好的结果，次高的数值代表稍差一点的结果。另外，在零和博弈的博弈表中，一个参与人的支付是越大越好，而另一个则是越小才越好。在这个囚徒困境中，越小的支付数值对博弈双方都是越好的。因此，如果要构造一个支付越大结果反而越差的博弈表，你就必须标明以提醒读者。同时，当你读到其他人的例子时，也要注意这样的问题。

下面我们从丈夫的角度来考虑表 4—4 中的囚徒困境博弈。丈夫必须考虑他的妻子如何选择。如果他认为妻子会选择认罪，那么他的最优选择也是认罪，这样就被判 10 年，因为若选择抵赖，他将被判 25 年。如果他认为妻子会要选择“抵赖”，情况又会怎样呢？他的最优选择仍是“认罪”，因为这样他仅被判 1 年，而选择抵赖则要被判 3 年。因此，在这个特殊的例子里，无论丈夫关于妻子行动的信念是什么，他选择“认罪”总是比选择“抵赖”好。我们称“认罪”是丈夫的**占优策略**（dominant strategy），或者称“抵赖”是**劣策略**（domi-

nated strategy)。相应地，我们可以称“认罪”占优于“抵赖”，或者称“抵赖”劣于“认罪”。

如果对于某个参与人而言，无论其他人如何选择，都有一个策略是最优的，那么我们就有足够的理由相信若此参与人是理性的，则必然会选择这个策略。另外，如果无论其他人如何选择，参与人的某个策略在所有的情况下都是比较差的，那么我们也可以肯定一个理性的参与人是不会选择这个策略的。因此，占优的存在为求解同时行动博弈提供了一个强有力的理论基础。

4.3.1 两个参与人都存在占优策略的情形

在上述囚徒困境的例子中，占优使丈夫选择了“认罪”。同理，对于妻子而言，“认罪”也占优于“抵赖”，所以她也将选择“认罪”。因此，(认罪，认罪) 构成了这个博弈可以预见的结果，并且这也是一个纳什均衡（事实上，也是该博弈唯一的纳什均衡），因为参与人都选择了他或她的最优策略。

在这个特殊的例子中，参与人的最优选择与他或她对对方行动的信念是否正确无关——这就是占优的含义，但是，由于对手也具有与自己相同的理性，所以每个参与人都必须具有正确的信念，并且其真实行动是对对手真实行动的最优反应。值得注意，对于两个参与人而言，“认罪”占优于“抵赖”的事实完全与他们到底是真的有罪（如多次在《纽约重案组》中发生的一样）还是被人陷害（如电影《幕后嫌疑犯》中发生的）无关，而只依赖于由不同入狱年限所决定的支付结构。

任何博弈若具有如表 4—4 所示的支付结构，就可以被称为是“囚徒困境”。更明确地说，囚徒困境有三个本质的特征：第一，每个参与人都有两个策略，与对手合作（在上述的例子中，就是“抵赖”）或“背叛”（“认罪”）；第二，每个参与人都有一个占优策略（“认罪”或“背叛”）；第三，对于两个参与人而言，占优均衡比选择劣策略时（即都选择抵赖或合作时）的非均衡要差。

学习博弈论，掌握囚徒困境这种类型的博弈是非常重要的，具体说来有两个原因：首先，类似囚徒困境的支付结构经常出现在经济、社会、政治，甚至生物竞争等诸多不同的策略场合中，广泛的应用性使得囚徒困境值得从策略的角度来加以理解和学习。整个第 11 章和其他章中的若干节将介绍这种博弈。

在讨论策略型博弈时，囚徒困境之所以重要的第二个原因是其均衡结果非常有趣。博弈双方都遵循常规的智慧选择了占优策略，但是这种均衡结果所产生的支付比他们都选择其劣策略时要低，所以对于参与人来说，囚徒困境的均衡结果是相对不好的。存在比均衡更好的博弈结果，但问题是如何确保参与人相互之间不会彼此欺骗。囚徒困境的这个特点引起了博弈论学者的广泛关注，他们提出了一个显而易见的问题：如何使囚徒困境中的参与人取得一个更好的结果？我们暂时把这个问题留给读者，并继续对同时行动博弈进行讨论，但我们将会在第 11 章中对这个问题做出详细的解答。

4.3.2 只有一个参与人存在占优策略的情形

当参与人存在占优策略时，他便会选择这个策略，并且其他人也相信他会这么做。在囚徒困境中，博弈双方都是如此。但在另一些博弈中，这种情况只适用于参与博弈的某一方。若你参与一个博弈，但你没有占优策略，而你的对手却有，那么你便会认为他要选择的就是他的占优策略，并且你也可以相应地选出均衡策略（即你的最优反应）。

我们用一个经常发生在国会和联邦储备银行之间的博弈[5]来说明这种情形。国会制定财政政策（税收和政府支出），而联邦储备银行负责货币政策（主要是利率）。简化这个博弈看其实质，国会的财政政策选择无非是“预算平衡”或“预算赤字”，联邦储备银行选择的是或高或低的利率水平。实际上，这一博弈并不能被清晰地判定为同时行动博弈，但同样也无从判定哪一方会率先采取行动，如果它们的行动存在先后顺序的话。因此，在这里，我们暂将这个博弈看成同时行动博弈。在第 6 章中，我们将看到当博弈规则不同时，博弈结果也会不同。

几乎所有的人都希望税负较轻，但对政府财政支持的要求却来自方方面面，如国防、教育、卫生环保等，同时还有各种具有政治影响的特殊利益团体（包括农民和被国际竞争损害的产业）需要得到政府的补贴。因此，国会一直都处于降低税收和扩大支出的双重压力下。然而，这些使预算产生赤字的举措可能会加重通货膨胀。联邦储备银行的任务是防止通货膨胀，但是它又不得不面临来自许多利益团体要求低利率的政治压力，特别是那些受益于低抵押贷款利率的私人房产主。较低的利率会刺激厂商对汽车、房屋和资产投资的更高需求，这些需求又会进一步引发高的通货膨胀。只要没有通货膨胀的威胁，联邦储备银行很乐意降低利率。而且，当政府预算平衡时，通货膨胀的威胁也相对较小。考虑以上的因素，我们构造了这个博弈的支付矩阵，如表 4—5 所示。

表 4—5　　货币政策—财政政策的博弈

		联邦储备银行	
		低利率	高利率
国会	预算平衡	3，4	1，3
	预算赤字	4，1	2，2

对于国会而言，“预算赤字”和“低利率”并存的结果是最好的（可以得到 4 的支付），并且也可以讨好政治中间派。就长远而言，这样的结果可能存在隐患，但每届国会的任期是短暂的，且其政治眼光也是短浅的。相反，国会最讨厌“预算平衡”和“高利率”并存的结果（只能得到支付 1）。在其他两个结果中，国会更青睐于“预算平衡”和“低利率”并存的结果（可以得到支付 3）。因为这样的结果可以得到那些拥有私人房产的中产阶级的支持，并且在利率较低时，削减政府债务就只需更少的财政支出，所以平衡预算为其他开支和减税创造了

空间。

对于联邦储备银行而言，“预算赤字”和“低利率”并存的结果是最坏的（只能得到支付1），因为这种策略组合将导致最高的通货膨胀风险。相反，它最喜欢“预算平衡”和“低利率”并存的结果（可以得到支付4），因为这样可以维持较高水平的经济活力，同时也没有过大的通货膨胀风险。在联邦储备银行选择“高利率”时，它更希望国会选择“预算平衡”，因为这样会降低通货膨胀风险。

我们下面来分析这个博弈中的占优策略。

如果联邦储备银行认为国会会选择“预算平衡”，那么它自己则会选择“低利率”（因为这样它可以得到支付4而不是支付3）；但如果它认为国会会选择“预算赤字”，则它选择“高利率”将会更好些（这样可以得到支付2而不是支付1）。由此可见，联邦储备银行不存在占优策略。然而，国会却有占优策略。如果国会认为联邦储备银行会选择“低利率”，国会的最优选择是“预算赤字”而不是“预算平衡”（这样可以得到支付4而不是支付3）；如果国会认为联邦储备银行会选择“高利率”，它的最优选择仍是“预算赤字”而不是“预算平衡”（这样可以得到支付2而不是支付1）。因此，“预算赤字”是国会的占优策略。

在这个博弈中，国会的选择是很清楚的，无论它认为联邦储备银行会如何选择，它都将选择“预算赤字”。联邦储备银行在自己决策时，也会考虑到这一点，它相信国会将选择其占优策略（预算赤字），并且在此信念上，联邦储备银行选择了自己的最优策略，即“高利率”。

在这个博弈的均衡结果中，双方的支付都是2。然而，考察表4—5，我们会发现，正如囚徒困境一样，存在另外一个可以带给博弈双方更高支付的结果，即（预算平衡，低利率）。其中，国会得到支付3，联邦储备银行得到支付4。为什么这个结果不能成为均衡呢？问题就在于国会存在偏离这个较好结果的动机，它会偷偷地选择预算赤字。联邦储备银行知道国会将这样做，并且这将给联邦储备银行带来最差的结果（只得到支付1），它也会偏离这个结果中的策略，而选择“高利率”。在第6章和第10章中，我们将考虑博弈双方如何走出这种困境，最终达成对于双方都更好的结果。但有一点值得说明，许多国家在许多时期，这两个政策的制定机构都陷入了坏的结果：财政政策过松，而为保持低的通货膨胀水平，货币政策往往过紧。

4.3.3 重复剔除劣策略

至今为止，我们所考虑的博弈中，每个参与人只有两个可选的纯策略。在这样的情形下，如果一个策略是占优策略，则另一个必然是劣策略，所以选择占优策略就等同于剔除劣策略。而在规模更大的博弈中，或许没有单个策略占优于其他所有的策略，但可能也存在一些策略劣于其他某些策略。如果参与人发现自己处于这样一种博弈中，他们可以通过将劣策略从可选策略中剔除的办法求得均衡。剔除劣策略缩小了博弈的规模，得到的“新”的博弈又可能出现对于同一个

参与人或其他参与人而言的另外一些劣策略，进而又可以继续剔除这些劣策略；或者在“新”的博弈中，某个参与人存在占优策略。**重复剔除劣策略**（successive or iterated elimination of dominated strategies）的过程可以剔除劣策略或缩小博弈的规模，直到不能进一步剔除为止。如果这个过程最后只剩下了一个唯一的结果，那么这个博弈就被称为**占优可解**（dominance solvable），这个结果就是博弈的纳什均衡，并且最终剩下的策略是每个参与人的均衡策略。

我们用表4—1中的博弈作为例子，来介绍重复剔除劣策略的过程。首先考虑“行”的策略，如果“行”的某个策略总是带来低于其他策略的支付，那么这个策略就是劣策略，并可将之从“行”的均衡策略中剔除。在这里，“高”是“行”唯一的劣策略，它劣于“底”。如果“列”选择“左”，“行”选择“底”得到支付5，而选择“高”仅得到4；如果“列”选择“中”，“行”选择“底”得到支付4，而选择“高”仅得到支付3；如果“列”选择“右”，“行”选择“底”得到支付9，而选择“高”仅得到支付6。所以我们可以剔除“高”。反过来，我们检查“列”是否存在可以被剔除的劣策略。我们发现“左”对“列”来说劣于“右”(推理同上，$1<2$，$2<3$，以及$6<7$)，但要注意，在“行”的“高”被剔除之前，我们不能这么说；因为若“行”选择“高”，“列”选“左”得到支付5，而选“右”仅得到支付4。因此，“行”的“高”被剔除是“列”的“左”被剔除的前提。在剩下的策略中（“行”还有“顶”、“低”、“底”，“列”还有“中”、“右”)，对于“行”而言，“顶”和“底”都劣于“低”。当“行”的策略只剩下“低”时，“列”便可选择其最优反应“中”。

因此，如表4—1所示的博弈是占优可解的，其结果是（低，中），支付是（5，4）。在原先利用这个例子解释纳什均衡的概念时，我们就已知这个结果是纳什均衡。下面我们来进一步考察参与人形成正确信念的思维过程。“行”若是理性的，就不会选择“高”；理性的“列”认识到这点，就会在“行”剩下的策略集中思考如何选择自己的策略，进而决定不选择“左”。反过来，“行”也认识到这些，进而不会选择“顶”和“底”。最终，“列”识破了所有的这一切，选择了“中”。

另外一些博弈不是占优可解的，或者重复剔除劣策略不能获得一个唯一的均衡。即使是这样，一些劣策略的剔除也可以缩小博弈的规模，并且使得采用后面小节中介绍的方法求解均衡更加容易。因此，即使当剔除劣策略不能完全解出博弈均衡的时候，劣策略的剔除也是解大型同时行动博弈的有用步骤。

至今为止，在我们所考虑的劣策略的剔除中，支付的大小都可以明确地做出比较。但是，如果某些支付相等，又该如何处理呢？将表4—3所示的博弈做一个变换，其中“高”(“行”的策略）和“左”(“列”的策略）也都被剔除掉。随后，“低”仍占优于“顶”，但“低”与“底”的占优关系就不那么明确了。

当“列”选“右”时，“行”选“低”比选“底”好；而当“列”选“中”时，“行”选“低”和“底”的支付相同。因此，对于“行”来说，我们称“低”弱占优于“底”。与此相对，“低”是强占优于“顶”的，因为在此时“列”只剩下两个策略“中”和“右”，“行”选“低”的支付要严格高于选“顶”的支付。

在本章附录中，我们将对强占优和弱占优的区别做严格的定义，而这里我们

仅提醒读者注意，重复剔除弱的劣策略可能会遗漏一些纳什均衡。

考虑另一个博弈，如表4—6所示。对“行”而言，“下”弱占优于“上”，因为当“列”选“左”时，“行”选“下”所得的支付比选“上”高；而当“列”选“右”时，“行”选“上”所得的支付和选“下”一样高。类似地，对“列”而言，“右”弱占优于“左”。占优可解告诉我们（下，右）构成纳什均衡，但我们还要注意到（下，左）和（上，右）也是纳什均衡。我们来看看（下，左）为什么是纳什均衡。当“行”选择“下”时，“列”不能通过背离“左”转向选择“右”而得到更高的支付；当“列”选择“左”时，“行”的最优反应是选择“下”。同理，我们可以验证（上，右）也是一个纳什均衡。

表4—6　　　　弱的劣策略的剔除

		列	
		左	右
行	上	0，0	1，1
	下	1，1	1，1

因此，如果你运用了弱占优来剔除某些策略，最好还得迅速逐个单元格检查一下，看看你是否错过了一些其他的均衡。重复剔除劣策略所得的占优解是对博弈中可能的纳什均衡的合理预测，但是仍然要注意多重均衡的存在性。在随后的章节中，我们将讨论这些问题，第5章将讨论多重均衡，第6章将讨论同时行动博弈和序贯行动博弈的相互联系。

4.4　最优反应分析

在许多同时行动博弈中，占优策略和劣策略是不存在的，或者尽管存在一个或几个劣策略，但劣策略的重复剔除不足以达到最终的唯一均衡。在这些情况下，我们需要借助其他的方法来进一步寻找博弈均衡。虽然我们要寻找的纳什均衡仍是每个参与人都在给定其他人选择的情况下做出的各自的最优选择，但我们将采用的方法是更加严密的策略思维推理，而不仅仅是简单的剔除劣策略。

下面我们来介绍一种系统地寻找纳什均衡的方法，你将在以后的分析中发现这种方法非常有用。开始，我们不要求信念的正确性，并轮流站在每个参与人的立场上思考以下的问题：在其他参与人可能的每种策略组合下，“我”的最优选择是什么？这样一来，我们就可以找出每个参与人对所有其他人可选策略的最优反应。用数学的语言来表达，就是每个参与人的最优反应策略依赖于其他人的可选策略，或者说是其他人可选策略的函数。

重新回到表4—1所示的博弈，并且将它复制成表4—7。首先考虑“行”。如果“列”选择“左”，“行”的最优反应是“底”，得到支付5。我们在这个博弈表中的相应支付上画圆圈来表示最优反应。如果“列”选择“中”，“行”的最优

反应是“低”（也是得到支付 5）。如果“列”选择“右”，“行”的最优反应仍然是“低”（可以得到支付 12）。和前面一样，我们在“行”的相应支付上画圆圈来标明其最优反应。类似地，我们在“列”的支付 3（“中”是对“行”选“顶”的最优反应）、支付 5（“左”是对“行”选“高”的最优反应）、支付 4（“中”是对“行”选“低”的最优反应），以及支付 7（“右”是对“行”选“底”的最优反应）上画圆圈，标明其最优反应。[6]我们发现单元格（低，中）中的两个支付都被画上了圆圈，因此就说明“行”的策略“低”和“列”的策略“中”同时是对对方策略的最优反应。这样，我们就找到了这个博弈的纳什均衡。

表 4—7　　　　最优反应分析

		列		
		左	中	右
行	顶	3，1	2，③	10，2
	高	4，⑤	3，0	6，4
	低	2，2	⑤，④	⑫，3
	底	⑤，6	4，5	9，⑦

最优反应分析（best-respond analysis）是一种全面寻找博弈中所有可能的纳什均衡的方法。你可以利用本章中的其他博弈来练习这种方法，进而提高对最优反应分析的理解。在存在占优策略的博弈中，运用这种方法是非常有意思的。如果“行”有占优策略，那么他对“列”的所有策略的最优反应都是一个相同策略，所以他的最优反应将水平地排列在同一行中。类似地，如果“列”有占优策略，那么他对“行”的所有策略的最优反应将会垂直地排列在同一列中。你可以运用这种方法，看看前面的囚徒困境和“国会－联邦储备银行”博弈的纳什均衡是如何出现的。

正如占优不能解出一些博弈的均衡一样，最优反应分析在一些博弈中也不能找到纳什均衡。虽然同样都是没有找到均衡，但最优反应分析却比占优分析提供了更多的信息。如果对一个离散型策略的博弈进行最优反应分析，没有发现纳什均衡，那么该博弈就不存在纯策略纳什均衡，我们将在第 8 章中分析这种类型的博弈。在第 5 章中，我们将最优反应分析运用到参与人策略是连续型变量的博弈中，例如价格或广告支出，并且通过绘制最优反应曲线来辅助寻找纳什均衡。由于策略选择的连续性，我们将会发现这种博弈几乎都有纳什均衡。

4.5　零和博弈的最小最大值方法

在零和博弈中，博弈双方的利益是完全对立的。基于这一特征，可以用**最小最大值方法**（minimax method）来替代最优反应分析方法。在零和博弈中，对其中一个参与人有利的结果就是对另一个参与人不利的结果。最小最大值方法就是

基于这一逻辑，且仅在零和博弈中适用。该方法假设每个参与人在选择策略时都会考虑："当其他参与人知道我所选择的策略时，这个策略还是我的最优选择吗?"因此参与人就得分析对手对自己策略的最优反应。然而，在零和博弈中，对手的最优反应行动对自己是最不利的。换言之，参与人都认为在自己所选行动的可能结果中，对手的选择会使其中最糟的结果出现。基于这样的信念，参与人就会选择使自己最坏的可能结果中的最好的行动。

这种逻辑看似非常消极，但仍是基于最优反应分析的，并且适用于在零和博弈中寻找均衡。在均衡中，每个参与人都会依据自己关于其他人行动的信念，选择最优反应。认识到这一点，参与人都清楚在自己所选的行动上只能得到最低的支付，从而以此选择自己相应的行动。因此，参与人会在这些最低的支付集合中选择一个带来最高支付的行动。

假设支付表仅列出了"行"的支付，那么"行"想要得到的结果就是单元格中数值尽可能地大，而"列"想要得到的结果却是单元格中的数值尽可能地小。"行"运用前述的悲观逻辑，他认为"列"将在他选择的每一行中选择数值最小的那一列策略。因此，"行"就要选择使得这些最小的数值最大的行策略，即在最小中选择最大——简而言之，就是**最大最小值**（maximin)。类似地，"列"也知道"行"将在他选择的每一列中选择数值最大的策略，那么"列"就要选择使得这些最大的数值最小的列策略，即在最大中选择最小——简而言之，就是**最小最大值**（minimax)。如果"行"的最大最小值和"列"的最小最大值出现在同一个单元格中，那么这个结果就是零和博弈的纳什均衡。这种寻找纳什均衡的方法应该被称为最大最小值－最小最大值方法，但简称为最小最大值方法。当纳什均衡存在时，你就可以运用这种方法找到纯策略纳什均衡。

我们借表 4—2 所示的橄榄球博弈来说明最小最大值方法。该博弈的纳什均衡，我们已经知道了，但不是通过最小最大值方法求得的。我们将博弈表重新复制为表 4—8，并且新添了一些适用于最小最大值方法的信息。

表 4—8　　　　最小最大值方法

		守方			
		跑动防守	传球防守	四分卫	
攻方	跑位	2	5	13	min=2
	短传	6	5.6	10.5	min=5.6
	中传	6	4.5	1	min=1
	长传	10	3	−2	min=−2
		max=10	max=5.6	max=13	

首先，我们找出每一行中的最小数值（即，攻方在自己实施每个策略时可得到的最低支付）和每一列中的最大数值（即，守方在自己实施每个策略时可得到的最低支付）。攻方选"跑位"时得到的最低支付是 2；选"短传"时得到的最低支付是 5.6；选"中传"时得到的最低支付是 1；选"长传"时得到的最低支

付为－2。我们将上述每行中的最小值标注在该行的右端。守方选择“跑动防守”时得到的最低支付是－10；选择“传球防守”时得到的最低支付是－5.6；选择“四分卫”时得到的最低支付是－13。[①] 我们将上述每列中的最大值标注在该列的下方。

下面我们就来寻找各个参与人可能得到的最差结果中最好的一个，即各行最小值的最大值和各列最大值的最小值。各行最小值的最大值是5.6，所以攻方将采用“短传”的策略以获得其最大最小值5.6码。各列最大值的最小值也是5.6，所以守方将采用“传球防守”的策略使得攻方仅仅得到5.6码，这也是守方的最小最大值。

观察这两个策略选择，我们发现最大最小值和最小最大值都出现在同一个单元格中。因此，攻方的最大最小值策略是对守方的最小最大值策略的最优反应，反之亦然。这样我们便找到了该博弈的纳什均衡。在均衡时，攻方试图短传，而守方将防止传球。最终，攻方将前进5.6码，同时守方也将退守5.6码。

在某些零和博弈中运用最小最大值方法，有时会找不到均衡。如果出现这样的情况，我们的结论和前面的最优反应分析是一致的：该博弈不存在纯策略纳什均衡。我们稍后将在本章中讨论这个问题，并且将在第7章和第8章中研究混合策略均衡。值得再次强调，不能在非零和博弈中运用最小最大值方法求解均衡。因为此时对手的最优选择并不一定使得你的情况最糟，所以符合悲观假设的令最小支付尽可能大的策略并不总是最优的。

4.6 三人参与的博弈

到目前为止，我们所讨论的博弈都只有两个参与人。然而，所有曾讨论过的寻找纯策略纳什均衡的分析方法都适用于包含任何数目参与人的同时行动博弈。若参与人的数目多于两人，且每人只拥有相对较少数目的纯策略，我们仍可以使用博弈表进行分析，正如在本章中的前5节一样。

在第3章中，我们曾描述过一个三人博弈，其中每个参与人只拥有两个纯策略，该博弈的三个参与人爱弥莉、莱娜和塔尼娅分别要决定是否为小街上的花园新建作贡献。假设三人都为花园作贡献不如仅有两人作贡献好；如果只有一人为花园作贡献，则花园中的花会太少，这样的结果和没有花园时一样糟。假设三人同时决策，且博弈的结果和支付有多种可能，花园的规模和美丽程度取决于贡献者的数量：如果三人同时都作贡献，花园最大和最美的；如果只有两人作贡献，花园的规模和美丽程度中等的；如果仅一人作贡献，花园是最小的，且最不美丽。

站在爱弥莉的角度考虑街心花园博弈，我们发现存在六种可能的结果。爱弥莉可以选择作贡献或不作贡献；与此同时，莱娜和塔尼娅都选择作贡献，或者都选择不作贡献，又或者她们中仅有一人选择作贡献。对于爱弥莉来说，当她的两

---

① 这些支付的相反数在表4—8中都是每列最大的正数。表4—8中的max和min分别表示“最大”和“最小”。——译者注

个好心邻居都选择作贡献而她自己选择不作贡献时，她将得到最好的结果，其支付为 6。此时，爱弥莉享有一个中等规模的花园，而不用自己花钱。如果爱弥莉同另外两人一样都选择了作贡献，她将享有一个大而美丽的花园，但由于自己也花了钱，所以在这种结果上她所得的效用是次高的，其支付为 5。

另一方面，当莱娜和塔尼娅都选择不作贡献时，爱弥莉也倾向于不作贡献。如果爱弥莉一人要承担大家共享的公共花园的所有费用，那么她还不如将花栽在自己的院子里。因此，若其他两人都不作贡献，爱弥莉认为自己选择作贡献时得到 1，而选择不作贡献时得到 2。

介于以上两种情况之间的是莱娜和塔尼娅中仅有一人为花园作贡献。此时，爱弥莉知道她可以在自己不作贡献的情况下享有一个小的花园。另外，她认为如果自己作贡献进一步扩大花园的规模，其成本超过了收益的增量。因此，她认为在自己选择不作贡献而享有一个小花园时，支付为 4；而自己作贡献将花园扩成中等规模时，支付为 3。关于作贡献和花园规模的成本和收益，莱娜和塔尼娅的观点与爱弥莉一致，所以她们也按爱弥莉的方法给不同的结果评分——如自己作贡献而其他人都不作贡献时是最差的，等等。

如果参与花园博弈的这三位女士在选择是否作贡献时不清楚相互的行动，那么该博弈就是一个三人参与的同时行动博弈。为了寻找该博弈的纳什均衡，我们需要列出支付表。对于三人参与的博弈来说，支付表应该是三维的，其中第三个参与人的策略对应于一个新的维度。要在二维支付表上增加一维，最简单的方法就是增加页面，支付表的第一个页面给出第三个参与人选择第一个策略时的各个参与人的支付，第二个页面给出第三个参与人选择第二个策略时的各个参与人的支付，依此类推。

街心花园博弈的三维支付表，如表 4—9 所示。行对应于爱弥莉的策略，列对应于莱娜的策略，页面则对应于塔尼娅的策略。我们将两个页面并排画在一起，以便于同时查看所有的结果。在各个单元格中，支付按行参与人、列参与人和页参与人的顺序排列，在本例中，依次便是爱弥莉、莱娜和塔尼娅的支付。

表 4—9　　　　街心花园博弈

塔尼娅作贡献

		莱娜	
		作贡献	不作贡献
爱弥莉	作贡献	5，5，5	3，6，3
	不作贡献	6，3，3	4，4，1

塔尼娅不作贡献

		莱娜	
		作贡献	不作贡献
爱弥莉	作贡献	3，3，6	1，4，4
	不作贡献	4，1，4	2，2，2

首先，我们考察各个参与人是否存在占优策略。在单页的支付表中，这种考察是非常简单的，我们只需比较参与人一个策略所对应的结果和她的另一个策略所对应的结果。对于行参与人而言，就是比较支付表上相对应的各行；对列参与人而言，就是比较相对应的各列。然而在这个三维支付表中考察是否存在占优策略，我们必须联合考虑支付表上的两个页面。

对于爱弥莉而言，比较支付表的两个页面上的各行发现：当塔尼娅作贡献时，爱弥莉的占优策略是不作贡献；当塔尼娅不作贡献时，爱弥莉的占优策略也是不作贡献。因此，无论其他人如何选择，爱弥莉的最优选择都是不作贡献。与此类似，我们还发现——在支付表的两个页面上——莱娜的占优策略都是不作贡献。当考察塔尼娅的占优策略时，我们就要小心一点。比较塔尼娅选择作贡献时和选择不作贡献时的支付，我们要保持爱弥莉和莱娜的行动不变，即我们要横跨支付表的两个页面来比较相对应的单元格——如比较第一个页面上位于左上角的单元格和第二个页面上位于左上角的单元格，如此等等。正如前两个参与人一样，分析得出不作贡献也是塔尼娅的占优策略。

在这个博弈中，每个参与人都有一个占优策略，这些策略的组合便构成了纯策略纳什均衡。在街心花园博弈的纳什均衡上，三个参与人都选择不作贡献，并且得到了次差的支付；花园没有建成，但没有人花了钱。

请注意，这个博弈又是一个囚徒困境的例子。在该博弈唯一的纳什均衡上，参与人所得的支付仅为 2，但该博弈的另一个结果——三个参与人都为花园作贡献——使得她们能得到更高的支付 5。虽然参与人都为建设花园作贡献对于她们来说是更好的，但是没有人有这样的动机。因此，花园以这种方式筹款是不可能被建成的，除非通过征税的方式来筹建，因为小镇的政府可以要求居民为此纳税。在第 12 章中，我们将碰到更多的这样的集体行动困境的例子，并将学习一些解决这类问题的方法。

寻找这个博弈的纳什均衡，也可以使用逐个单元格检查的方法。例如，考虑表 4—9 中的另一个单元格，其中爱弥莉和莱娜都选择作贡献而塔尼娅选择不作贡献，各人所得支付为（3，3，6）。当爱弥莉考虑改变策略时，作为行参与人的她只能通过换行来改变博弈的结果，即从一个位于某行、某列、某页的单元格移动到一个不同行但相同列、相同页的单元格。如果爱弥莉从支付为（3，3，6）的单元格出发改变策略，她的支付将从 3 提高到 4。与此类似，莱娜可以通过换列将结果移动到不同列但相同行、相同页的另一个单元格。在本例中，莱娜若改变策略，她的支付将从 3 提高到 4。最后，塔尼娅也可以通过换页且保持行和列不变来改变博弈的结果。在本例中，塔尼娅这样做会使自己的支付从 6 降低到 5。因为至少有一个参与人可以通过单边地改变策略来使自己的情况变好，所以支付为（3，3，6）的单元格不是纳什均衡的结果。

我们也可以用最优反应分析来求解这个博弈的均衡。我们可以和表 4—7 一样，在最优反应的支付上画上圆圈，如表 4—10 所示。因为不作贡献对于每个参与人来说都是占优策略，所以爱弥莉的最优反应都是在不作贡献那一行上，莱娜的最优反应都是在不作贡献的那一列上，塔尼娅的最优反应也都是在不作贡献那一页上。最终，我们找出了纳什均衡。

表 4—10　运用最优反应分析解街心花园博弈

塔尼娅作贡献

		莱娜	
		作贡献	不作贡献
爱弥莉	作贡献	5，5，5	3，⑥，3
	不作贡献	⑥，3，3	④，④，1

塔尼娅不作贡献

		莱娜	
		作贡献	不作贡献
爱弥莉	作贡献	3，3，⑥	1，④，④
	不作贡献	④，1，④	②，②，②

4.7　纯策略博弈的多重均衡

到目前为止，我们所考虑的博弈都只有一个纯策略纳什均衡，但一般说来，博弈可能包含有多个均衡。我们将用一类应用较广的博弈来说明这种情况。这类博弈被称为**协调博弈**（coordination games），其中参与人之间存在某些共同（但不完全一致）的利益，但是由于他们的决策是相互独立的（非合作博弈的性质），要协调行动以达到共同偏好的结果也不是那么容易。

为了更好地说明上述观点，设想有两个大学生哈里和萨莉，他们在学校的图书馆相遇，并且聊得很愉快。正当意犹未尽之时，他们又马上得回到各自的教室去上课，所以他们约好在 4：30 下课后一起喝咖啡。可是正当上课的时候，他们惊奇地发现刚才忘了约定见面的地方。有两个地方可供选择，星巴克和本地咖啡店，但不巧的是，这两个地方分别位于偌大的校园的两个相反方位，所以不能同时去这两个地方。另外，哈里和萨莉又没有交换过手机号码，故又不能发信息联系。那他们该怎么办呢？

将这个事件看成一个博弈，并用支付矩阵来刻画，如表 4—11 所示。每个参与人有两个选择——星巴克和本地咖啡店，如果他们相遇，则所得支付为 1；如果没有相遇，则支付为 0。运用逐个单元格检查的方法，我们发现这个博弈有两个纳什均衡，一个是他们都选择去星巴克，而另一个是他们都选择去本地咖啡店。达成均衡对于他们双方来说都是有利的，但具体是哪一个均衡得到实现就无关紧要了，因为他们在这两个均衡上的支付相等。采取哪一个行动并不重要，关键是他们的行动要协调一致，这就是这种类型的博弈被称为**完全协调博弈**（pure coordination game）的原因。

表 4—11　　完全协调博弈

		萨莉	
		星巴克	本地咖啡店
哈里	星巴克	1，1	0，0
	本地咖啡店	0，0	1，1

然而，他们能够行动统一吗？或者是否最终他们分别到了不同的咖啡店，还认为是对方爽约了？后一种可能是存在的。哈里可能认为萨莉会去星巴克，因为萨莉曾说过她上课的教室在校园中星巴克的那一边。但萨莉想的可能与哈里的行动完全相反。当多个纳什均衡存在时，参与人若想要成功地选择，他们就需要用某些方法来统一关于对方行动的信念和期望。

上述情况与第 1 章中“哪个轮胎爆了”的博弈相似，在那里我们将协调机制称为**聚点**（focal point）。在本例中，可能两家咖啡店中的某一家是学生们众所周知的活动场所，但是仅仅哈里知道这一点是不够的，他还得确信萨莉也知道，以及萨莉知道他也知道，依此类推。换言之，他们的期望必须收敛于聚点。否则，哈里不能确定萨莉要去哪里，因为他不知道萨莉对他要去哪里是怎么想的，相同的疑问将会出现在第三阶、第四阶或更高阶数的期望上。

我们中的一位（迪克西特）曾在课堂上向学生提出过这个问题，大一的学生一般会选择星巴克，大三和大四的学生往往会选择校园学生活动中心的咖啡店。这样的回答是非常容易理解的——大一的学生由于来校园的时间较短，所以他们的期望都集中在大家熟知的全国连锁的咖啡店；在大三和大四的学生眼中，适应当地的消费习惯是更好的选择，所以他们会期待他们的同伴也这么想。

如果某家咖啡店用橙色装饰，而另一家用深红色装饰，那么在普林斯顿，前者就是一个聚点，因为橙色是普林斯顿的标志色；而在以深红为标志色的哈佛，后一家咖啡店才是一个聚点。如果一个是普林斯顿的学生而另一个是哈佛的学生，那么他们就很难相遇了，这是由于他们都认为自己的标志色应该成为首选，或者都认为对方很顽固而试图迁就对方。一般说来，协调博弈中的参与人要找到聚点，就必须要有某些他们都知道的“连接点”，无论是历史的、文化的，还是语言的。

现在我们稍微改动一下该博弈的支付。大三和大四学生的行为表明博弈双方对去哪一家咖啡店并不是一点都不在意，可能某家的咖啡比另一家好，或者某家的氛围比另一家好，或者他们想去一家学生不常去的咖啡店，以免碰到以前的男朋友或女朋友。假设博弈双方都喜欢本地咖啡店，则他们在本地咖啡店相遇时的支付为 2，而当在星巴克相遇时的支付为 1。新的支付矩阵如表 4—12 所示。

表 4—12　　安全版本

		萨莉	
		星巴克	本地咖啡店
哈里	星巴克	1，1	0，0
	本地咖啡店	0，0	2，2

在这个新的博弈中，仍有两个纳什均衡，但参与人更偏好双方都去本地咖啡店的均衡。然而，喜欢这个结果并不代表能实现这个结果，重要的（也是我们经常分析到的）是支付必须是共同知识——博弈双方必须知道整个支付矩阵，并且知道对方也知道，如此等等。如果博弈双方曾讨论过这两家咖啡馆的相对优点，并且达成了共识，在博弈中仅仅是忘了明确地说在本地咖啡店会面，那么他们就具备了这样的共同知识。即便这样，哈里也可能认为萨莉有理由选择星巴克，或者认为萨莉对他的行动也会这么想，如此等等。如果关于行动的预期没有形成真正的**预期收敛**（convergence of expectations），他们有可能会走向较差的均衡，更有甚者，他们不能统一行动，都只得到为 0 的支付。

再次强调，在如表 4—12 所示的博弈中，参与人仅当对对方选择适当的行动有充分信心时，他们才能达成双方都更偏好的均衡。因此，这个博弈又被称为**安全博弈**（assurance games）。

在现实生活中，如果参与人之间进行一定的沟通，上述信心是容易具备的。当参与人之间的利益完全一致时，如果一人对另一人说："我要去本地咖啡店"，则另一人是不会怀疑这句话的真实性的，并且他也会去那里，从而得到相互都满意的结果。这也是为什么我们要用两个分别在不同地方上课且无法沟通的学生来编造这个故事的原因。如果参与人之间的利益是冲突的，开诚布公的沟通就会有问题，我们将在第 9 章中讨论信息的策略性操纵时分析这个问题。

在更大规模的团体中，沟通可以通过安排会议或发布消息来实现。这些方式仅当每个人都知道其他所有人都在关注这些方式的时候才有效，因为成功协调是使渴望得到的均衡成为一个聚点的前提，所有参与人的期望必须收敛于这个聚点，并且每个人都知道其他人也知道……同时，每个人都选择这个聚点上相应的策略。许多社会制度和规则就起到了这样的作用。开会时，参会者通常是围坐成一个面朝里面的圆圈，目的就是为了让每个人都知道其他人也在关注。"超级碗"（Super Bowl）期间的广告就是为了让每位观众都知道许多其他人和自己一同在观看，特别是当它们被提前宣布以吸引注意时。

基于这个道理，如果某些商品的消费者看到其他人购买时自己就更想买，那么生产这些商品的厂商就非常喜欢做广告，其中包括电脑业、通讯业和互联网业。[7]

下面我们再在咖啡店选择博弈中加入新的元素，设想两个参与人渴望相遇，但他们分别喜欢的是不同的咖啡馆。若在星巴克会面，哈里可以得到为 2 的支付，而萨莉只能得到为 1 的支付；若在本地咖啡店 会面，萨莉可以得到为 2 的支付，而哈里只能得到为 1 的支付。这个支付矩阵如表 4—13 所示。

表 4—13　　　　性别战

		萨莉	
		星巴克	本地咖啡店
哈里	星巴克	2，1	0，0
	本地咖啡店	0，0	1，2

该博弈被称为**性别战**（battle of the sexes），这个名字源于 20 世纪 50 年代的女权主义运动中的一个故事，是博弈论学者为说明这种支付结构而特地编造的。故事讲的是，丈夫和妻子要选择到底是去看拳击比赛还是去看芭蕾舞，（可能是由于进化中的遗传因素）丈夫喜欢看拳击比赛，而妻子则喜欢看芭蕾舞。“性别战”这个名字就由此得来，我们在这里也继续使用它，但要声明我们的这个例子没有性别歧视色彩。

这个博弈的结果如何呢？答案是仍然存在两个纳什均衡。如果哈里认为萨莉会选择去星巴克，那么他的最佳选择也是去星巴克，反过来也是如此。同样的道理，本地咖啡店也是一个纳什均衡。为了实现其中某一个纳什均衡，并避免两个人去不同咖啡店的结果，参与人需要一个聚点或者预期收敛，就好像在完全协调博弈和安全博弈中一样。但在性别战中，协调失败的风险要更大一些。参与人最初的情况是完全对称，但在两个纳什均衡上，他们的支付是不对称的，而且他们对这两个均衡的偏好是有冲突的，哈里更喜欢在星巴克会面，而萨莉更喜欢在本地咖啡店会面。他们必须要找到一种打破对称的办法。

为了达到他或她所偏好的均衡，参与人可以试着表现得强硬一点，从而引导对方选择相应的策略去实现那个对自己有利的均衡。这种参与人采用的达成自己所偏好均衡的事先措施被称为策略性行动，我们将在第 10 章中详细讨论。另外，参与人也可以表现得和善一点。比如上一次萨莉为了让哈里高兴，陪他一起去了星巴克，那么这一次哈里为了回报萨莉，他便会选择去本地咖啡店，虽然这对自己是相对不利的。在欧·亨利（O. Henry）的短篇小说《麦琪的礼物》（*The Gift of the Magi*）中，夫妇为对方选择圣诞礼物也有一个与此相类似的故事。再者，如果博弈是可重复的，协调也可通过谈判来成功地实现。例如，参与人可以事先商定轮流去这两家咖啡店。在第 11 章关于囚徒困境的讨论中，我们将分析这种发生在重复博弈中的默许合作。

本节中的最后一个例子也是一个协调博弈，但与前面的例子稍有不同。在该博弈中，参与人要尽量避免而不是选择那些一致的行动。此外，协调失败的后果也是非常悲惨的。

这个例子源于发生在 20 世纪 50 年代的死亡游戏，两个美国青年开着车于午夜时分在美国某地（Middle-of-Nowhere）的主干道（Main Street）上相向而驶，他们试图撞向对方。如果谁调头回避相撞，那么这个人就是“懦夫”；如果谁勇敢径直向前，那么这个人就是“勇士”。可以想象，两人都选择径直向前的后果是车毁人亡。[8]

懦夫博弈（chicken）的支付取决于参与人对“坏”结果的评价和认可程度，即是车毁人亡更“坏”还是被认为是懦夫更“坏”。如果参与人认为尊严受损要好于车毁人亡的话，如表 4—14 所示的支付结构就是对 20 世纪 50 年代的懦夫博弈的合理刻画。每个参与人都希望获胜，让对方成为懦夫，同时也最不希望发生车祸。除此之外，双方都成为懦夫的结果要比只有自己当懦夫的结果好。

表 4—14　　　　　　　懦夫博弈

		迪恩	
		转向（懦夫）	径直向前（勇士）
詹姆斯	转向（懦夫）	0，0	−1，1
	径直向前（勇士）	1，−1	−2，−2

懦夫博弈有四个本质的特征：第一，每个参与人都同时有一个"强硬"策略和一个"软弱"策略；第二，博弈有两个纯策略纳什均衡，其中在每个均衡上恰好有一个参与人是懦夫或表现得软弱；第三，每个参与人都严格偏好对手选择当懦夫或表现得软弱的均衡；第四，两个参与人都表现得强硬的结果是最差的。在这类博弈中，真正的博弈变成了双方就如何达到各自所偏好的均衡展开的较量。

我们现在回到了与曾讨论过的"性别战"相似的情形，可以想象真实生活中的懦夫博弈要远比"性别战"残酷——获胜的收益很大，但相撞的损失也很大。因此，利益的冲突和参与人支付的不对称性等问题都在这里被强化了。每个参与人都试图影响博弈的结果。一种可能的情形是造成大家都认为他必定采取强硬的策略的氛围以威胁对手。[9]另一种可能情形则是通过做出显而易见的和无可置疑的要径直向前的承诺，让对手相信自己是不会当懦夫的（在第 10 章中，我们将讨论如何做出这样的承诺行动）。此外，如果可能的话，博弈双方都会试图避免悲惨（撞车）结果的发生。

我们曾在"性别战"中谈过，如果博弈被重复多次，默许的协同是达成均衡的较好途径。这里也是如此，如果这两个青年在每周六的午夜都来玩这个游戏，则他们在考虑均衡策略时便会参考过去和未来的结果。这样一来，他们就会理性地轮流选择这两个均衡，也就是轮流当勇士和懦夫（但是，如果还有其他人知道这个规则的话，博弈双方则会颜面无存）。

最后，在这些协调博弈中有个问题值得注意，纳什均衡的概念要求每个参与人都具有关于其他人策略选择的准确信念。当我们寻找纯策略纳什均衡的时候，纳什均衡的概念要求每个参与人都确信其他人的选择。然而，在我们关于协调博弈的分析中，参与人对其他人行动的期望充满了不确定性，那么又如何将这些不确定性纳入我们的分析呢？在第 5 章中，我们将引入混合策略的概念，即真实的选择是可选行动的随机组合，有了这样的理解，纳什均衡的概念就可推广到参与人不能确信相互行动的情形。

4.8　不存在纯策略纳什均衡的情形

到目前为止，我们所考虑的每个博弈都至少有一个纯策略纳什均衡，如第 7 节中的那些博弈有多个均衡，在更前面的章节中某些博弈则仅有一个均衡。然而，在学习策略和博弈论时，并非我们所遇到的所有博弈都存在这种可简单加以

定义的解——其中，每个参与人总是选择某一特定的行动作为均衡策略。在本节中，我们将讨论那些不存在纯策略纳什均衡的博弈——在这些博弈中，没有参与人一直都选择某一个特定行动来作为均衡策略。

网球比赛是个简单的例子，可以用来说明某些博弈没有纯策略纳什均衡的存在。设想一场比赛，由两个顶尖的女子选手纳芙拉蒂洛娃和埃弗特对决。在网前的纳芙拉蒂洛娃把球打给站在底线的埃弗特，而埃弗特要准备回击，她可以选择打底线球（DL；快速直球）或者是打斜线球（CC；慢速对角球）。纳芙拉蒂洛娃也必须准备往哪边防守，每个参赛者都清楚她们的行动计划不能让对手知道，因为这会被对手利用。纳芙拉蒂洛娃可以朝埃弗特将会击球的那边防守，而埃弗特可以把球打到纳芙拉蒂洛娃没有防守的那一边。如果我们假设双方都能隐藏自己的意图，直到要行动的最后一刻，那么她们的行动就是同时进行的，这场网球赛也是同时行动博弈。

网球博弈中的支付，是指在某一特定击球和防守的回合中双方赢球的概率。假设底线球比斜线球更有威胁力；当纳芙拉蒂洛娃防守时选错边，埃弗特更有可能获胜。这样我们可以求得合理的支付结构：如果纳芙拉蒂洛娃防守斜线球，则埃弗特打底线球有 80%的获胜概率；如果纳芙拉蒂洛娃防守底线球，则埃弗特打底线球只有 50%的获胜概率；同理，如果纳芙拉蒂洛娃防守底线球，则埃弗特打斜线球有 90%的获胜概率；如果纳芙拉蒂洛娃防守斜线球，则埃弗特打斜线球的获胜概率降到 20%。

显然地，纳芙拉蒂洛娃赢球的概率是 100%减去埃弗特赢球的概率，所以这是一个零和博弈（更准确地说，这是常和博弈，因为双方的支付之和为 100%）。我们仅需在支付表中写出埃弗特的支付即可，表 4—15 表示四种不同可能结果时埃弗特赢球的概率。

表 4—15　　纯策略均衡不存在的情形

		纳芙拉蒂洛娃	
		DL	CC
埃弗特	DL	50	80
	CC	90	20

解同时行动博弈的规则告诉我们，首先是找占优策略或劣策略，然后再尝试用最小最大值方法或逐个单元格检查的方法来寻找纳什均衡。可以发现，双方均不存在占优策略，那么就继续用逐个单元格检查的方法。从双方都选底线球开始，埃弗特可以采用改打斜线球将获胜概率从 50%提高到 90%，然后纳芙拉蒂洛娃又可以采用改防守斜线球将埃弗特的获胜概率降低到 20%，之后，埃弗特可以选择底线球又将获胜概率提高到 80%，随后，纳芙拉蒂洛娃又会防守底线球。所以在博弈表的每一个单元格中，都会有一个参与人想改变策略，然后我们就沿着博弈表无休止地绕圈，找不到任何均衡。

在这类不存在纯策略纳什均衡的博弈中，关键不是参与人应该做什么，而是参与人不应该做什么。尤其是在面对这种情形时，参与人不应该总是选择相同的

策略，否则对手将会有机可乘（如果埃弗特总是打斜线球，纳芙拉蒂洛娃就会察觉到这点，并且每次都防守斜线球，从而降低了埃弗特赢球的概率）。参与人理性的出招方式不应该服从某种规律，而应力求出奇制胜，所以每一种策略的采用都只能持续一段时间（埃弗特应该以足够的频率使用她比较弱的策略，这样纳芙拉蒂洛娃才猜不准她的击球方式，但是她也不应该以固定形式采取这两种策略，因为这样的战术也可为对方所掌握）。如果参与人随机选择她的行动，那么她使用的策略就被称为混合策略（mixing strategy），这在第 7 章和第 8 章中将被着重讨论。表 4—15 所示的博弈不存在纯策略纳什均衡，但却存在混合策略纳什均衡，第 7 章的第 1 节将告诉我们这一结论是如何得到的。

4.9 总　结

在同时行动博弈中，参与人必须在不知道其他人行动的情况下做出选择。这种博弈可用博弈表来描述，博弈表的各个单元格给出了参与人的支付，博弈表的维度等于参与人的数目。在二人零和博弈中，博弈表仅需给出其中一个参与人的支付即可。

纳什均衡是同时行动博弈中解的概念，它是一个由参与人策略所构成的组合，其中每个参与人只选择一个策略，每个策略是对其他人策略的最优反应。纳什均衡也可被定义为这样一种策略组合：其中每个参与人都具有关于其他人策略选择的准确信念，并且给定这些信念，参与人的策略都是最优的。寻找纳什均衡，可以采用逐个单元格检查法、寻找占优策略法、重复剔除劣策略法以及最优反应分析等。零和博弈还可以用最小最大值方法来求解。

本章还介绍了多种类型的同时行动博弈。囚徒困境是经常遇到的一种；协调博弈包括安全博弈、懦夫博弈和性别战等，一般存在多个均衡，达成这些均衡需要参与人以某种方式协调行动。虽然某些博弈不存在纯策略纳什均衡，但可能存在混合策略纳什均衡，关于混合策略纳什均衡的详细讨论请见第 7 章和第 8 章。

关键术语

安全博弈（assurance game）
性别战（battle of the sexes）
信念（belief）
最优反应（best response）
最优反应分析（best response analysis）
逐个单元格检查（cell-by-cell inspection）
懦夫博弈（chicken）
常和博弈（constant-sum game）
预期收敛（convergence of expectations）
协调博弈（coordination game）
占优可解（dominance solvable）
占优策略（dominant strategy）
劣策略（dominated strategy）
枚举法（enumeration）
聚点（focal point）
博弈矩阵（game matrix）

博弈表（game table）
最大最小值（maximin）
最小最大值（minimax）
最小最大值方法（minimax method）
混合策略（mixed strategy）
纳什均衡（Nash equilibrium）
规范式（normal form）
支付表（payoff table）
囚徒困境（prisoners' dilemma）
完全协调博弈（pure coordination game）
纯策略（pure strategy）
重复剔除劣策略（successive or iterated elimination of dominated strategies）
策略式（strategic form）
零和博弈（zero-sum game）

习　　题

1. “在同时行动博弈中，如果某个参与人有一个占优策略，那么他就可以得到对于他来说最好的结果。”请问这句话对不对？解释你的答案，并举出一个博弈的实例来说明。
2. 采用检查是否有占优策略和重复占优的方法，寻找下列零和博弈的纯策略纳什均衡，然后用最小最大值方法加以验证。

(a)

		列	
		左	右
行	上	1	4
	下	2	3

(b)

		列	
		左	右
行	上	1	2
	下	4	3

(c)

		列		
		左	中	右
行	上	5	3	1
	直	6	2	1
	下	1	0	0

3. 找出下列非零和博弈的所有纯策略纳什均衡，并给出你的分析步骤。

(a)

		列	
		左	右
行	上	2，4	1，0
	下	6，5	4，2

(b)

		列	
		左	右
行	上	1，1	0，1
	下	1，0	1，1

(c)

		列		
		左	中	右
行	上	0，1	9，0	2，3
	直	5，9	7，3	1，7
	下	7，5	10，10	3，5

(d)

		列		
		西	中	东
行	北	2，3	8，2	10，6
	上	3，0	4，5	6，4
	下	5，4	6，1	2，5
	南	4，5	2，3	5，2

4. 验证下面的博弈是否是占优可解的，并且找出纳什均衡。

		列		
		左	中	右
行	上	4，3	2，7	0，4
	下	5，5	5，−1	−4，−2

5. 找出下列博弈的所有纯策略纳什均衡，描述你寻找的过程，并且借助这个博弈解释为什么要使用参与人的策略而不只是均衡时的支付来描述均衡。

		列		
		左边	中间	右边
行	上面	1，2	2，1	1，0
	水平	0，5	1，2	7，4
	下面	−1，1	3，0	5，2

6. 俾斯麦海［以将俾斯麦群岛（Bismarck Archipelago）和巴布亚新几内亚（Papua-New Guinea）分隔开来的太平洋西南海域命名］战役是第二次世界大战期间美国和日本之间的一场海战。1943 年，日本的舰队司令得到命令将护卫舰队转移到巴布亚新几内亚，但他必须选择是沿一条多雨的北部航线还是沿一条阳光灿烂的南部航线转移，这两个方案都要 3 天的航行时间。美军知道日军的舰队要转移并试图轰炸它，但他们不知道日军将取道哪一条航线。美军决定派出侦察机去打探日军的航线，但是现有的侦察机一次只够去一条航线上执行任务。因此，日军和美军都是在不知道对方计划的情况下决定各自的行动。

   如果日军的舰队正好在美军侦察到的航线上，美军将马上轰炸它；如果不是这样，美军则延误一天轰炸的时间。北部航线恶劣的天气不利于轰炸，如果美军在北部航线上正好发现了日军，它可以得到 2 天的轰炸时间；如果美军到北部航线上侦察而日军的舰队却沿南部航线航行，美军仍可以得到 2 天的轰炸时间。此外，如果美军开始到南部航线上侦察，正好发现日军则得到 3 天的轰炸时间，而没有发现则只能得到 1 天的轰炸时间。

   (a) 用博弈表描述这个博弈。

   (b) 找出这个博弈中的占优策略，并且解出纳什均衡。

7. 有一个老妇人在过马路时需要得到帮助，只要有一个人帮助她就可以了，很多人帮助她当然好，但不会比一个人帮助她时好到哪里去。你和我在这位老妇人附近，都可以提供帮助，并且我们都同时决定是否帮助她。无论我们谁帮助她，老妇人顺利过了马路就可以让我们都得到价值为 3 的快乐。但是施以帮助的人要为自己的善举付出价值为 1 的时间成本。请依据这个故事建立一个博弈，写出它的支付表，并且找出它的所有纯策略纳什均衡。

8. 两个参与人杰克（Jack）和吉尔（Jill）被隔离在不同的房间里，同时他们被告知一个博弈的规则，每个人从 G、K、L、Q、R、W 这六个字母中选择一个字母，如果他们选择一致，则可以得到以下奖赏：

字母	G	K	L	Q	R	W
杰克的奖赏	3	2	6	3	4	5
吉尔的奖赏	6	5	4	3	2	1

   如果他们选择了不同的字母，则均得到 0。整个规则都被告诉给了两个参与人，并且双方都被告知对方也知道这个规则，如此等等。

   (a) 画出这个博弈的支付表，纯策略纳什均衡有哪些？

   (b) 均衡中有没有一个是聚点？哪一个是？为什么？

9. 假设两个参与人 A 和 B，选择三个不同的策略 1、2、3，如果他们选择相同，就会得到奖金，如以下博弈表所示。

		B		
		1	2	3
A	1	10，10	0，0	0，0
	2	0，0	15，15	0，0
	3	0，0	0，0	15，15

(a) 找出博弈的纳什均衡，哪一个有可能是聚点？请解释。

(b) 我们把博弈改变一下，两个支付为（15，15）的单元格变为（25，25），其余的不变。如果双方掷硬币来决定采取策略 2 或策略 3，则参与人的期望（平均）支付是多少？是否比以（1，1）为聚点的支付还高？如果 A 和 B 采取不同的方式决定采取何种策略，你如何评价这种风险？

10. 在第 3 章中，爱弥莉、莱娜和塔尼娅进行了一场序贯的街心花园博弈，其结果有四种情况（而不像在 4.6 节的例子中有六种结果），对于每个参与人来说，这四种结果是：

(1) 参与人自己不作贡献，而另外两个参与人都作贡献（建成了美丽舒适的花园，并且节约了自己作贡献的成本）

(2) 参与人自己作贡献，但另外两人也都作贡献或只有一人作贡献（建成了美丽舒适的花园，但付出了作贡献的成本）

(3) 参与人自己不作贡献，其他两人中只有一人作贡献或他们都不作贡献（花园花草稀疏，但节约了自己作贡献的成本）

(4) 参与人自己作贡献，但另外两人都不作贡献（花园花草稀疏，同时自己付出了作贡献的成本）

在以上四种均衡中，第一种对该参与人来说是最好的（支付为 4），第四种是最差的（支付为 1）。如果每个参与人认为花园的美丽舒适比自己作贡献付出的成本要重要，那么她将在第二种结果中得到 3，在第三种结果中得到 2。

(a) 假设这三个参与人在不知其他人选择的情况下，同时决定是否为花园作贡献。画出这个三人博弈的博弈表。

(b) 找出博弈中的所有纳什均衡。

(c) 设想在现实中这种同时行动的花园博弈是如何进行的。

11. 考虑一个博弈，它能提供价值为 30 美元的奖品。有三个竞争者 A、B、C 参与，每个人可以选择购买价值为 15 美元或 30 美元的票，或根本就不买票，他们独立地同时做出决策。在知道所有人的买票决策后，博弈的组织者将按如下规则颁发奖品：如果没有人买票，那么奖品就不颁发给任何人；反之，如果有人买票，那么奖品就颁发给购买的门票价值最高的人。如果购买的门票价值最高的只有一人，那么奖品就属于他一人；如果有多人购买的门票同时价值最高，那么奖品就在他们间平分。给出这个博弈的策略式，并且找出所有纯策略纳什均衡。

12. 在电影《美丽心灵》(*A Beautiful Mind*) 中，约翰·纳什 (John Nash) 和他的 3 个研究生同学在酒吧里遇到了一个困难的抉择，他们可以去和 4 位浅黑肤色女生和 1 位金发女生搭讪，但是他们每个人只能接近 1 位女生并吸引她的注意。如果吸引到金发女生，能得到 10 的支付；如果吸引到浅黑肤色女生，能得到 5 的支付；如果没有吸引到任何女生，则支付为 0。那么如何才能吸引到这些女孩呢？规则是如果两个或多个小伙子走向金发女生，那么他们都会被金发女生拒绝，并且也会被那些浅黑肤色女生拒绝，因为那些女生不想当小伙子们的第二选择。由此可见，每个参与人仅当他是唯一一个接近金发女生的人时才会得到 10 的支付。

(a) 首先考虑只有 2 个小伙子而不是 4 个小伙子的简单情形。(相应地，只有 2 个浅黑肤色女生和 1 个金发女生，但她们对小伙子们的反应如前所述不变，并且她们不是博弈的参与人。) 画出博弈的支付表，找出所有的纯策略纳什均衡。

(b) 当有 3 个小伙子的时候 (同时有 3 个浅黑肤色女生和 1 个金发女生，并且她们仍不是参与人)，画出这个 (三维的) 博弈表，并找出所有的纳什均衡。

(c) 将在 (a) 题和 (b) 题中的分析推广到小伙子的人数为 4 的情形，进一步扩展到小伙子人数为任意数 n 的情形。无须画出 n 维的支付表，仅需分析当 k 个参与人走向金发女生而 $(n-k-1)$ 个选择接近浅黑肤色女生时参与人的支付，其中 $k=0,1,\cdots,(n-1)$。电影中作为均衡的结果——所有的小伙子都选择了浅黑肤色女生——是这个博弈的真实纳什均衡吗？

【注释】

[1] 事实上，价格只能以货币制度的最小单位来度量 (例如美分)，所以价格也是一个有限离散型取值的变量。但是这样的单位太小以至于将价格看成连续变量更为合理一些。

[2] 这里我们来解释该支付是如何计算出来的。当攻方选择中传，同时守方选择传球防守，我们估计有 50%的概率传球成功并得到 15 码，40%的概率传球失败 (得到 0 码)，10%的概率传球被中途拦截，得到 −30 码。这样一来，攻方的期望所得为 $0.5\times15+0.4\times0+0.1\times(-30)=4.5$ 码。表中的这些数值是由迪克西特召集的具有专业知识的邻居和朋友组成的小组计算出来的，他们得到了流动的咨询费 (意为不固定的咨询费——译者注)。

[3] 这个概念以数学家和经济学家约翰·纳什的名字命名。纳什于 1949 年在普林斯顿大学获得博士学位，他还提出了合作博弈中的一个解的概念 (我们将在第 17 章中介绍)。他与其他两位博弈论大师莱茵哈德·泽尔腾和约翰·海萨尼 (Reinhard Selten and John Harsanyi，我们将在第 9 章和第 10 章中介绍他们的一些成就) 一起分享了 1994 年的诺贝尔经济学奖。西尔维亚·娜莎 (Sylvia Nasar) 为纳什写的传记《美丽心灵》(*A Beautiful Mind*, New York: Simon & Schuster, 1998) 是由罗素·克鲁 (Russell Crowe) 主演的一部电影的主要根据。但是这部电影对纳什均衡的解释是错误的。我们将在本章的习题 12 和第 8 章的习题 9 中解释这个败笔。

[4] 但是，注意支付为（5，5）的策略组合（底，中）不是一个纳什均衡。若“行”选择“底”，“列”的最优选择不是“中”，他可以通过选择“右”而变得更好。事实上，你可以检查其他的单元格来验证没有其他的纳什均衡。

[5] 类似的博弈发生在许多中央银行能独立制定和实施货币政策的国家。但在不同的国家，财政政策可能是由不同的政治机构（行政机构或立法机构）制定的。

[6] 等价地，我们也可以在那些没有选择的策略的相应支付上做记号。如在表 4—3 中，“行”没有将“顶”、“高”、“底”选为对“列”的策略“右”的最优反应，我们就可以用斜线划去“行”的相应支付（这里就是 10、6、9）来标明。按如此方法标记完博弈双方的所有策略以后，发现（低，中）中的两个相应策略都没有被划去，所以这就是博弈的纳什均衡。在最优反应的策略上画圆圈和把绝非最优反应的策略划去实质上是一致的，就如同在序贯行动博弈中，可以采用在被选的树枝上画箭头或“剪去”那些没有被选的树枝两种方法。我们更喜欢使用前一种方法，因为这样可以使最终的图形和求解博弈的过程更清楚。

[7] Michael Chwe 在“*Rational Ritual*：*Culture*，*Coordination*，*and Common Knowledge*”（Princeton：Princeton University Press，2001）中对这个问题有深入的介绍。

[8] 懦夫博弈有一个不同的版本，出现于 1955 年的电影《阿飞正传》（*Rebel Without a Cause*）中。有两个人平行飞快驱车冲向断崖，在到达断崖前，首先跳车的人是懦夫，太晚跳的人会冒着坠崖身亡的危险。影片中把这称为“胆怯游戏”（chicky-game）。20 世纪 60 年代中期，英国哲学家鲁塞尔（Bertrand Russell）和其他和平主义者用懦夫博弈比喻美苏之间的核武竞赛，博弈论学者拉波特（Anatole Rappoport）提供了规范的博弈论分析，其他学者认为核武竞赛是囚徒困境或安全博弈。请参考 Barry O'neill，“Game Theory Models of Peace and War”，in *Handbook of Game Theory*，Vol. 2，ed. Robert J. Aumann and Sergiu Hart，Amsterdam：North Holland，1994，pp. 995 - 1053。

[9] 如果有一个人，他从未在懦夫博弈中退让过，则谁会跟他玩呢？问题在于，懦夫博弈中，参与人可能并不是自愿的，如在诉讼中一样。换个角度来说，选择是否参与懦夫博弈本身就是一个懦夫博弈。谢林说过：“如果你被邀请参加懦夫博弈，而你拒绝了，则你就已经参加并且输了”（*Arms and Influence*，Yale University Press，1965，p. 118）。

附录　几个一般性的定义

我们在本章中曾通过一些例子介绍了占优和纳什均衡的概念，虽然这些可能对你们来说已经够用了，但了解更为准确的定义仍是有益的。我们考虑参与人为"行"和"列"的一般化博弈，假设"行"有四个可选的策略，分别被记为 R1、R2、R3 和 R4，"列"有三个策略，分别是 C1、C2 和 C3。有这些策略就足以表达必要的意思了，更多的策略也不能揭示什么额外的含义。

这个"行""列"博弈的支付表如表 4A—1 所示，在这个一般化设定中，我们不使用具体的数值来表示支付，而是采用一般的代数符号。表示每个支付的代数符号由一个字母及两个数字组成，其中字母代表是谁的支付，其后的第一个数字表明"行"的策略，第二个数字表明"列"的策略。如在图中的阴影单元格中，"行"选择了 R3，"列"选择了 C2，那么"行"的支付就是 R32，"列"的支付是 C32。

表 4A—1　　一个一般性的博弈

		列		
		C1	C2	C3
行	R1	R11，C11	R12，C12	R13，C13
	R2	R21，C21	R22，C22	R23，C23
	R3	R31，C31	R32，C32	R33，C33
	R4	R41，C41	R42，C42	R43，C43

假定（R3，C2）这个策略组合就是纳什均衡，那么它的支付应满足什么条件呢？在给定“行”关于“列”选择的信念是 C2 时，R3 是“行”的最优选择。因此，当“列”选择 C2 时，R32 是“行”可选择的不同策略所得到的支付中最高的，即

R32≥R12，R32≥R22，R32≥R42

注意在上面的定义中，我们用了大于等于号“≥”而不是严格的不等号“>”，所以我们允许了相等的情况存在。类似地，给定“列”关于“行”选择的信念是 R3 的情况下，“列”的最优选择 C2 应满足

C32≥C31，C32≥C33

当我们检验一个特定的单元格是不是纳什均衡时，我们要验证所有的这些不等式，即“行”在这个单元格中的支付应是该列中其所有支付中最大的（或这个支付至少不低于其他支付），同时“列”在这个单元格中的支付应是该行中其所有支付中最大的（或这个支付至少不低于其他支付）。

我们继续用这个博弈表来定义占优。假定对“列”而言，C2 占优于 C1，那么无论“行”选择什么，“列”选择 C2 的支付总是比选择 C1 时高。我们进一步区别下面两种情况：如果 C2 上的每个支付都严格大于 C1 上的支付，我们就称 C2 强占优于 C1，即

C12>C11，C22>C21，C32>C31，C42>C41

如果 C2 上的每个支付至少不小于 C1 上的支付，我们就称 C2 弱占优于 C1。但至少在某一行中，C2 上的支付严格大于 C1 上的支付，否则 C2 和 C1 上的支付就完全相等了。此外，至少在某一行中，等式成立，否则又成强占优了，因此，

C12≥C11，C22≥C21，C32≥C31，C42≥C41；其中至少有一个“=”和一个“>”成立

例如，在表 4—1 的博弈中，对“行”而言，“底”强占优于“高”（因为 5>4，4>3 和 9>6）。但在表 4—3 中，在“列”的策略“左”被剔除后，“低”仅仅弱占优于“底”（因为 5=5，12>9）。

第 5 章　纯策略同时行动博弈（Ⅱ）：连续型策略、（Ⅲ）：讨论和证据

在第 4 章中，关于同时行动博弈的讨论主要集中于参与人的可选策略是离散型集合的情形，同时我们还介绍了一些离散型策略同时行动博弈的例子，其中包括在某些情况下招数不多且容易定义的体育比赛，比如在足球比赛中罚点球的时候，射手的可选策略是往高处射或是往低处射，往边角射或是朝正中央射。另外，我们还介绍了协调博弈和囚徒困境，在这些博弈中，每个参与人仅拥有两个或三个可选策略。博弈表可以用来辅助分析离散型策略同时行动博弈，特别是当博弈中参与人数目和可选策略数目都不太大的时候。

然而，许多同时行动博弈与我们曾讨论过的例子不同，参与人策略选择的区间可能是一个广阔的范围，供参与人选择的策略可能有无限多个，比如制造企业为它们的产品定的价格、慈善家捐款的数额，以及竞标者确定的投标金额等。客观上来讲，价格和以货币度量的金额都是有最小单位的，如美分，所以可选的价格策略也是一个有限的和离散的集合。但是由于这些单位太小，在实际运用时往往将这些单位计量的策略看成连续变化的实数。当参与人拥有如此多的可选策略时，博弈表仅能作为思考的工具，运用博弈表分析就显得不切实际了。就连续型策略同时行动博弈而言，我们需要一种新的求解方法。本章的第一部分将介绍处理**连续型策略**（continuous strategy）博弈的分析工具。

本章还会就与纳什均衡的概念和同时行动博弈中的行为有关的若干问题展开广泛的讨论。我们将回顾所获得的或源于实验或来自现实生活的关于纳什均衡的经验证据。我们还将介绍一些对纳什均衡概念的理论批评和对这些批评的反驳，

你将认识到基于博弈论的预测往往是理解真实行为的合理起点。另外，你还将得到一些告诫，比如在参与人的层次上。

5.1 纯策略是连续型变量的情形

在第 4 章中，我们运用了最优反应分析的方法来寻找同时行动博弈中的纯策略纳什均衡，现在我们把这种方法的应用扩展到参与人可选策略是连续区间的情形，例如，公司为其产品进行定价。为了计算这类博弈中的最优反应，我们要分别在一家公司的所有可选价格下计算另一家公司（最大化其支付）的最优定价。由于策略集的连续性，我们可以将这些最优反应在图中用曲线画出来，图中的坐标轴线表示相应的参与人的定价策略（或其他连续型策略），纳什均衡就是两条曲线的交点。我们举两个例子来说明这种思想和方法。

5.1.1 价格竞争

我们的第一个例子发生在一个名为雅皮港（Yuppie Haven）的小镇，镇上有两家餐馆——泽维尔的达帕斯西班牙餐馆（Xavier's Tapas Bar）和伊冯娜的炭火西餐馆（Yvonne's Bistro）。为了简化这个例子，我们假设餐馆按菜单收费，并且它们分别设置各自菜单上的价格。因此，价格是餐馆竞争博弈中的策略。餐馆经营的目标是设置合理的价格以求利润的最大化，也就是追求博弈中支付的最大化。我们还假设餐馆在菜单上定价的时候是对对方的价格一无所知的，则这个例子是同时行动博弈。[1]因为价格可以取（几乎）无限区间中的任意值，所以我们就暂且用一般化的代数符号来表示价格，其中泽维尔的价格记为 P_x，伊冯娜的价格记为 P_y。我们将使用**最优反应规则**（best-response rules）来解这个博弈并找到它的均衡价格。

每家餐馆在定价的时候，必须考虑随之产生的利润。为了简化问题，我们假设两家餐馆的成本和收益情况是完全对称的，但数学基础好的读者可以使用更一般的数据或者代数符号来做类似的分析。假设每家餐馆为每个顾客提供服务的成本是 8 美元。根据经验或市场调查，若泽维尔的定价为 P_x，同时伊冯娜的定价为 P_y，则它们各自顾客的数量 Q_x 和 Q_y（单位是百人每月）由以下方程给出[2]：

$$Q_x = 44 - 2P_x + P_y$$
$$Q_y = 44 - 2P_y + P_x$$

从上面两个方程的形式可以看出，如果一家餐馆将价格抬高 1 美元（比如说伊冯娜将 P_y 抬高 1 美元），那么它的顾客每月将会减少 200 人（Q_y 的减少量是价格的 2 倍），而另一家餐馆的顾客每月将增加 100 人（Q_x 的增加量是价格的 1 倍）。可以猜测，伊冯娜的 100 位顾客转到泽维尔去消费，还有 100 位就留在家里用

餐了。

泽维尔每周的利润 B_x（B 表示"账本底线"）就是它从每位顾客身上得到的净收益（价格减去成本，即 P_x-8）乘以它所接待的顾客数量：

$$B_x=(P_x-8)Q_x=(P_x-8)(44-2P_x+P_y)$$

泽维尔的最优反应规则是：在伊冯娜的每个可能价格 P_y 下，设置自己相应的价格 P_x，以最大化自己的利润。我们稍后将画出最优反应曲线。

在每个 P_y 下计算使得 B_x 最大化的价格 P_x，有两种方法。第一种方法就是整理前面 B_x 表达式的右边各项：

$$\begin{aligned}B_x&=-8(44+P_y)+(16+44+P_y)P_x-2(P_x)^2\\&=-8(44+P_y)+2(15+0.25P_y)^2-2(15+0.25P_y-P_x)^2\end{aligned}$$

在上面式子的第二行中，只有最后一项含有 P_x，并且这一项的前面是减号，那么要使得整个式子（即泽维尔的利润）最大化，就要求这一项尽可能地小。同时，这一项又是一个平方式，所以它的最小值就是 0，则该项为 0 时满足下式：

$$P_x=15+0.25P_y$$

上式给出了在伊冯娜的每一个价格 P_y 下使得泽维尔的利润最大化的价格 P_x。换句话说，这就是我们想要得到的泽维尔的最优反应规则。

另外一种求解最优反应规则的方法是运用微分。在前面 B_x 表达式的第一行中，对 P_x 求导数（保持另一家餐馆的价格 P_y 不变）：

$$\frac{dB_x}{dP_x}=(60+P_y)-4P_x$$

使得 B_x 最大化的一阶条件就是上面这个导数等于 0。令上式等于 0 并解出 P_x，所得结果与第一种求解最优反应规则方法所得的结果一样（B_x 的二阶导数为 -4，满足二阶条件 $d^2B_x/dP_x^2<0$）。

伊冯娜的最优反应规则也可以使用上述两种方法求得。由于两家餐馆成本和销量完全对称，伊冯娜的最优反应规则为

$$P_y=15+0.25P_x$$

两家餐馆的最优反应规则同样也可以用最优反应曲线来表示。例如，如果泽维尔将价格定为 16，那么伊冯娜就将泽维尔的价格代入自己的最优反应规则，计算得出 $P_y=15+0.25(16)=19$；同样地，泽维尔对伊冯娜的价格 $P_y=16$ 的最优反应是 $P_x=19$。各家餐馆对对方价格为 4 的最优反应都是 16；对对方价格为 8 的最优反应都是 17，依此类推。

这两个最优反应规则，如图 5—1 所示。由于我们例子的特殊性——销售数量和价格的关系是线性的，生产每份食品的成本为常数，所以这两条**最优反应曲线**（best-response curve）都是直线。如果改变关于需求和成本的定义，最优反应曲线就不一定是直线了，但是求解它们的方法是不变的：首先令一家餐馆的价格（如 P_y）保持不变，计算使得另一家餐馆利润最大化的价格（如 P_x），再反过来用同样的方法计算 P_y。

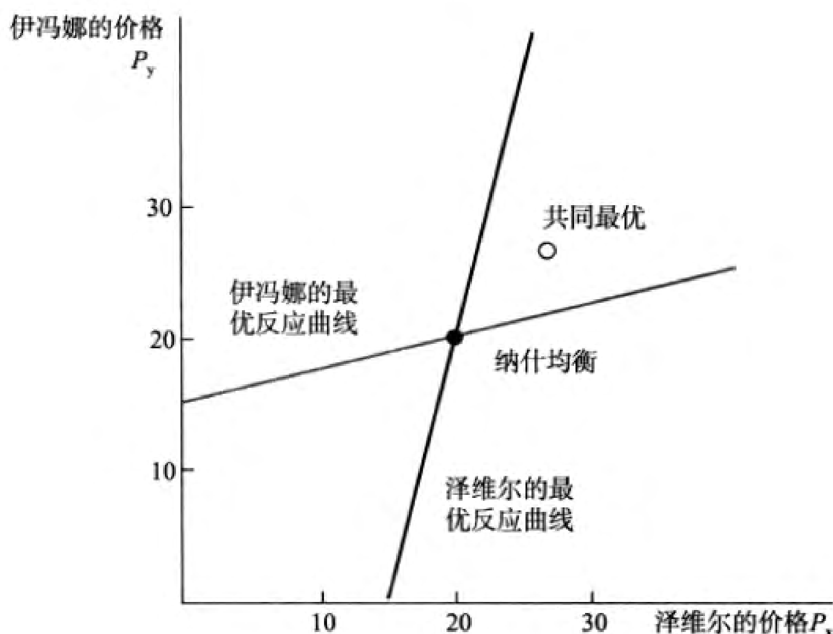


图 5—1　餐馆定价博弈的最优反应曲线和均衡

两条最优反应曲线的交点就是餐馆定价博弈的纳什均衡，它代表了一组价格，其中每家餐馆价格都是对另一家价格的最优反应。餐馆的均衡价格 P_x 和 P_y 的具体取值可以通过联立求解两个最优反应规则而得到，由于方程是线性的，所以求解也很容易。我们将 P_x 的表达式带入 P_y 的表达式，便可得到下式：

$$P_y = 15+0.25P_x = 15+0.25(15+0.25P_y)$$
$$= 18.75+0.0625P_y$$

将上式化简，求出 $P_y=20$。基于解的对称性，很容易得到 $P_x=20$。[3]因此，在均衡时，每家餐馆的定价均为 20 美元，它们每月服务的顾客人数都是2 400人[2 400=(44−2×20+20)×100]，并且从每位顾客身上赚取的利润是 12 美元，所以它们每月的利润总额均为 28 800 美元。

介绍本例的主要目的是学习如何在连续型策略（如价格）的博弈中求解纳什均衡，但是在这个例子中关于厂商定价和利润的经济学道理是非常有趣的，值得进一步的讨论。

首先，我们发现每条最优反应曲线都是向上倾斜的。如果一家餐馆将价格提高 1 美元，另一家的最优反应就是将自己的价格提高 0.25 美元或 25 美分。当一家餐馆提价时，它的一些顾客就转到另一家餐馆消费，并且另一家餐馆为了最大化其利润，也部分地提高了价格。因此，一家餐馆的提价行为有助于增进另一家餐馆的利润。在纳什均衡上，各家餐馆的定价是相互独立的，所以它们在定价时都只考虑自身利润，而没有考虑由于自身行为转移到对手的利润。那么，它们能够通过合作提价来增进双方的利润吗？答案是肯定的。如果两家餐馆都将价格提高到 24 美元，则每家餐馆每月将服务 2 000 位顾客 [2 000=(44−2×24+24)×100]，并且从每位顾客身上赚取 16 美元的利润，所以它们各自的利润将提高到 32 000 美元。

这个定价博弈与我们在第 4 章中介绍的囚徒困境非常相似，只不过这里的策略是连续型变量。在第 4 章的那个例子中，夫妇二人都有动机向警方招供，并背叛对方。然而，如果他们都这么做，双方都将被判更长的刑期（一个更糟的结果）。与此类似，虽然将价格定为 24 美元能得到更高的利润，但这不是纳什均衡，两家餐馆在独立决策时会削减这个价格。如果伊冯娜开始将价格定为 24 美元，运用最优反应规则，我们可以知道泽维尔会将价格定为 $15+0.25\times24=21$ 美元，再后来，伊冯娜的最优反应将是 $15+0.25\times21=20.25$ 美元。继续这个过程，双方的价格最终将收敛于均衡价格 20 美元。

然而，什么样的价格是对两家餐馆来说的共同最优价格呢？考虑到对称性，我们假设双方的定价都为 P，那么各家的利润为

$$B_x=B_y=(P-8)(44-2P+P)=(P-8)(44-P)$$

它们都选择能使得上式极大化的 P 值，运用前面讲过的任意一种求极值的方法，我们很容易求出 $P=26$，此时每家餐馆每月的利润分别为 32 400 美元。

用经济学的术语来说，这种联合将价格提高到共同最优水平的合谋称为卡特尔（cartel）。高价损害了消费者的利益，所以美国政府的监管机构常常打击卡特尔，以促进市场竞争。公开的价格合谋是非法的，但仍有可能保持默许的合谋，正如在重复囚徒困境博弈中可能发生的那样——我们将在第 11 章中讨论这种情况。[4]

合谋并不一定总是导致更高的价格。在前面的例子中，如果某家餐馆降价，它的销量会上升，部分原因是由于两家餐馆所提供的产品互为替代品，通过降价行为可以从对手那里吸引到更多的顾客。在其他情况下，两家公司可能销售的是互补的商品——如硬件和软件。那么，如果其中一家公司降价，两家公司的销量都会上升。在纳什均衡上，各家公司的定价是相互独立的，并且没有考虑到降低自己产品的价格可以增进对方的利润。因此，它们将价格定得比合作时高。如果允许它们之间合作，那么价格就会更低一些，并且这样的结果对消费者来说也是更有利的。

5.1.2 政党宣传拉票

第二个例子源于政治，其中运用的数学会比我们平时用的多一些，但我们同时还将用文字解说和图表来帮助大家理解这些计算背后的含义。

有两个党派或两个候选人参加一场竞选，他们都试图通过宣传的方式来为竞选拉票——可能是正面的宣传，借以强调自身的优势；也可能是负面的、攻击性的宣传，借以强调对手不光彩的一面。为了简化例子，我们假设选民们起初都是完全无知的和漠不关心的，并且他们的决策仅仅受到这些宣传的影响（许多人都认为这就是美国政治的真实写照，但在政治科学中更为严谨的分析承认某些选民是见多识广的和具有策略眼光的，我们将在第 15 章中详细探讨这些选民的行为）。进一步简化，我们假设每个党派所获得的选票份额就等于他在竞选宣传中

所占的份额。如，参加竞选的党派或候选人分别被称为L和R，若L为宣传花费了 x 百万美元，同时R为宣传花费了 y 百万美元，那么L将得到的选票份额为 $x/(x+y)$，R将得到的选票份额为 $y/(x+y)$。筹集宣传资金也是有成本的，为简化分析，我们假设这些成本与直接的竞选支出成比例。特别地，我们令党派L的支付为 $x/(x+y)-x$；类似地，党派R的支付是 $y/(x+y)-y$。

下面我们来分析党派的最优反应。由于在求解中不能不使用微积分，所以我们先用数学导出表达式，然后再用文字语言解释其中的含义。给定党派L的策略 x，党派R将选择 y 以最大化其支付。那么，求解中就要求 x 保持不变，令 $y/(x+y)-y$ 对 y 求导的值为0就得到了一阶微分条件，即 $x/(x+y)^2-1=0$，或 $y=\sqrt{x}-x$。同理，党派L的最优反应方程是 $x=\sqrt{y}-y$。方程的曲线图如图5—2所示。

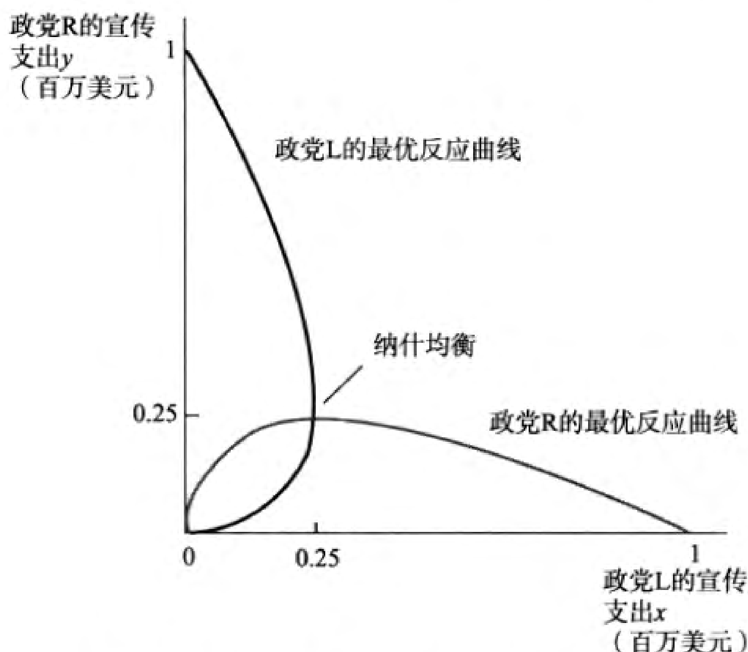


图5—2　竞选宣传博弈中的最优反应和纳什均衡

让我们来看看党派R的最优反应曲线。当党派L的宣传费用增加时，党派R的宣传费用先是短暂地增加然后转为下降。如果对手宣传的费用很少，那么自己为宣传投入的费用就有一个很高的选票回报，并且当对手小幅增加宣传费用时，自己也会以增加宣传费用作为回应。然而，如果对手已经为宣传花费了很多，那么自己宣传的收益相对于成本来说是微不足道的，所以在这个时候回应对手增加宣传费用的更好方式是自己减少宣传费用。

非常巧合，这两个党派的最优反应曲线在它们的顶点处相交。我们用代数方法解这两条曲线所代表的联立方程组，便得出在均衡上的 x 值和 y 值均等于1/4，即250 000美元（可以想象，这场竞选是小型的地方性竞选），读者自己也可以加以验证。

与餐馆定价博弈一样，这个博弈又是一个囚徒困境。若两个党派都同比例地削减宣传费用，它们各自的选票份额将不受影响，但双方都因节省了开支而提高

了所得的支付。生产替代品的厂商之间形成的卡特尔使得市场价格较高并且损害了消费者的利益；而生产互补品的厂商之间形成的卡特尔则导致了市场低价。政治家之间形成的卡特尔更像后者而不是前者，它们将减少宣传费用，进而有可能提高选民和整个社会的福利。因此，找到解决这种囚徒困境的办法对大家来说都是有益的。事实上，国会已经为此付出了多年的努力，并且取得了部分成效，但是政治竞争非常激烈，至今还没有一个方案是完美的。

如果两个党派间的实力不对等，结果又会如何呢？有两种可能导致实力不对等。一是可能由于某个党派（如 R）和媒体的关系更好，它能够以较低的成本进行宣传；二是可能党派 R 的宣传花费比党派 L 更加有效，比如 L 的选票份额是 $x/(x+2y)$，而 R 的份额是 $2y/(x+2y)$。

如果是上述的第一种情况，R 会充分利用自己更便利的宣传途径，因此它会在任何 L 的宣传费用 x 上都选择一个比原来更高的费用 y。那么，图 5—2 中的党派 R 的最优反应曲线将向上移动，纳什均衡也沿着保持不变的党派 L 的最优反应曲线向西北方向移动。因此，最终党派 R 的宣传费用会比党派 L 的多，处于优势的党派会利用自己的优势，而处于劣势的党派则会甘拜下风。

在第二种情况下，两个党派的最优反应曲线的移动就比较复杂了。结果还是双方都投入一样多的宣传费用，但都少于它们在实力对等时的宣传费用 x。如在我们的例子中，党派 R 的宣传费用的效率是党派 L 的两倍，则最后双方的宣传费用都是 x（由此可见，实力对等时的竞争最为激烈）。当党派 R 的花费更有效时，双方的最优反应曲线是不对称的，新的纳什均衡不是在两条最优反应曲线的顶端，而是在 L 最优反应曲线的下降部分和 R 最优反应曲线的上升部分。换言之，虽然博弈双方最终在宣传上的花费相同，但处于优势的党派 R 的宣传花费要比导致党派 L 最大反应时的花费多，而处于劣势的党派 L 的宣传花费要比导致党派 R 最大反应时的花费少。在本章的一个选做习题（习题 10）中，数学好的同学可以自己推导出这个结论。

5.2 纳什均衡的经验证据

在第 3 章中讨论序贯行动博弈和逆向归纳的经验证据时，我们曾说过证据可以源于对现实世界中实际发生的博弈的观察，也可以来自为检验理论而在实验室或教室里人为设计的博弈，我们也曾指出了这两种方法各自的优缺点。在寻找和解释同时行动博弈的证据时，也会遇到同样的问题。

真实世界的博弈是由那些富有经验的参与人参与的，并且他们在博弈中都有实质性的利益，这使得他们都有知识和动机去选择正确的策略。但是这些环境中还含有许多理论未顾及的因素，所以如果数据不能证实理论的预测，我们也就不能判定到底是理论错了还是一些其他的因素左右了策略的选择。

实验室实验可以控制其他因素，因而可以提供“干净”的测试，但是参与这些实验的参与人是没有经验的，并要在较短的时间和较少的激励下学习和开展博

弈。面对一个新的博弈，许多人都会彷徨失措，随机尝试，所以实验室环境中的博弈的头几轮只代表了这种学习的过程而不是富有经验的参与人所要选择的均衡。一些实验考虑到参与人的经验不足和学习过程，常常将头几轮的数据舍弃。但是学习的过程可能不止一个上午或一个下午，所以这也是实验室实验的主要弊端。

5.2.1 实验室或课堂里的实验

在过去的30年间，研究者进行了不计其数的实验室实验，用以验证当人们处于某些特定的策略互动场合时他们是如何反应的，特别是考察他们是否选择了纳什均衡的策略。道格拉斯·D·戴维斯（Douglas D. Davis）和查尔斯·A·霍尔特（Charles A. Holt）[5]通过回顾这些研究发现：如果博弈相对简单，只进行一次且只有一个纳什均衡，那么“在与不同的参与人重复地进行博弈后……理论有很强的预测能力”；但如果博弈是复杂的，或是重复的，或是由于存在多个均衡而需要某种协同才能达到均衡，或是需要通过计算才能找到复杂的纳什均衡，那么基于理论的预测就不一定成功了。我们将简要地讨论在这些情形下纳什均衡的表现。

1. 在多个均衡中选择

当博弈存在多个均衡时，参与人一般都不能很好地协调，除非他们具有达成聚点均衡的共同文化背景（也就是共同知识）。托马斯·谢林（Thomas Schelling）和戴维·克雷普斯（David Kreps）报道了一些协调博弈的试验。[6]克雷普斯让他的学生成对地玩如下的博弈：一个学生被指派波士顿，另一个被指派旧金山。给他们每个人都发了一张写有美国其他九个城市的名单——亚特兰大、芝加哥、达拉斯、丹佛、休斯敦、旧金山、纽约、费城以及西雅图——并且让他们从中选择一部分城市。他们的选择是同时进行的且相互独立的。如果他们的选择正好涵盖了这九个城市且没有重叠，双方都可得到奖励；否则，他们都得不到任何奖励。这个博弈有多个（512个）纯策略纳什均衡，但是如果参与这个博弈的参与人都是美国人或长期的美国居民，他们将会有超过80%的可能按地域划分；被指派波士顿的学生会选择密西西比河以东的所有城市，被指派旧金山的学生会选择密西西比河以西的所有城市。然而，如果两个学生都不是美国居民，这种协调就很难发生，此时他们往往都是按字母顺序选择，但是这种选择没有明确的划分界限。

2. 天生利他主义的显示或实验中的集体主义精神

某些实验环境中的参与人行为与实验设计者预期的纳什均衡不符，其中一个方面就是人们常常“错误”地走到友好和公平的一边。因此，我们观察到在囚徒困境博弈中有“太多”的合作，在讨价还价博弈中人们“太”迁就对方，但导致这些现象的原因可能不是参与人在计算或学习选择纳什均衡过程中的错误，而可能是参与人的支付与实验设计者设想的不太一样。

和在真实世界中发生的博弈一样，实验环境中的参与人可能比实验设计者本

身更了解环境的复杂性，比如博弈重复的可能性，或当时某种与其他参与人的关系会影响他们在博弈中的决策。另外，参与人的价值体系中可能包含了友好和公平等社会价值，这些价值在社会环境中被广泛地认同，因而也影响了人们在实验中的行为。[7]这些现象并不是说明纳什均衡概念本身无效，而是提醒我们不要在关于参与人支付的天真和错误的假设下运用纳什均衡的概念。比如，参与人总是为对金钱的追求所驱动这一假设，可能是错误的。

3. 从经验中学习

有一个游戏，经常在课堂上或实验室里玩，就是叫每个参与人在 0 到 100 之间选择一个数，然后交上写有他们名字和选择数字的卡片。由此可见，这个游戏是一个同时行动博弈。当卡片收上来以后，计算所有卡片上数字的平均数，选择的数字最接近平均数的某个比例——比如说一半——的人将是赢家。这个博弈的规则（游戏的整个过程）在事先就被宣布。

该博弈的纳什均衡是每个人都选择 0。事实上，这个博弈是占优可解的，即便所有人都选择 100，平均数的一半也不会超过 50，所以对于每个参与人来说任何高于 50 的选择都劣于 50。[8]但是，全部参与人都会出于理性想到这一点，所以平均数不会超过 50，这个数的一半就不会超过 25，那么所有高于 25 的选择都劣于 25。依此类推，相继占优，直到最终收敛到 0。

然而，当一组人第一次玩这个博弈时，赢家选择的通常是比 25 稍小的数字。该结果意味着，赢家假设其他人的选择都是随机的（所以平均数是 50），然后在此基础上选择了自己的最优反应。这个结果是远远偏离纳什均衡的。

如果这个博弈被重复进行多次，结果又将如何呢？根据在课堂上实验的经验，我们发现赢家选择的数字在各轮博弈中呈迅速下降的趋势。在斯克丝的班上，首先一半的学生玩这个博弈，同时另一半学生观看；然后，反过来，刚才观看的学生玩这个博弈；最后全班都玩这个博弈。在迪克西特的班上，每次 10 个不同的学生一组玩这个博弈，到了第三轮，赢家的选择通常只是 2 或 3。

如何解释这个结果呢？批评纳什均衡的观点认为，除非纳什均衡恰好达到，否则，这个理论就该被推翻。他们认为，如果你有足够的理由相信别人不会选择纳什均衡上的策略，那么你的最优策略也不是纳什均衡上的策略；如果你猜测到别人的选择将会如何偏离纳什均衡策略，那么你将选择针对他们策略的最优反应。但是，一些人则反驳，社会科学中的理论是不可能像物理学和化学一样做到精确的预测的，如果观察到的结果足够逼近纳什均衡，那这就是对理论的一种支持。在这种情况下，实验不仅提供了这种支持，而且还揭示出了人们积累经验进而选择逼近纳什均衡策略的过程。我们也是赞同后一种观点的。

有趣的是，我们发现人们在观察别人博弈时比自己参与博弈时学习得更快，这可能是因为身为旁观者可以关注博弈的全局，并自由地思考和分析；而作为参与人，思维则集中于为自己的决策做分析，缺乏广阔的视角。

我们下面来解释什么是“在博弈中积累经验”。本节一开始就引用了戴维斯和霍尔特的话“与不同的参与人重复地进行博弈”，换言之，就是在经常与不同的对手进行博弈的过程中积累经验。然而，对于任何使得结果逼近纳什均衡的学习过程，全部学习者的人数要保持稳定，如果新手不断地出现并尝试新的策略，

则原来的人就可能忘记在和其他人博弈中积累的经验。

如果博弈在两个参与人之间或在一小群熟悉的人之间重复进行，则会经常遇到相同的对手。在这种情况下，重复的博弈整体来看也是一个博弈，它的纳什均衡可能不是单次博弈纳什均衡的简单重复。比如，在重复的囚徒困境博弈中可能出现默许的合作，原因是由于大家都清楚背叛的短暂得利不如随即失去信任后的损失大。如果博弈如此重复进行，那么学习这种博弈，就要经常进行整套的重复博弈，并且每次的对手都不同。

5.2.2 现实世界的证据

基于博弈论的预测有两种途径接受来自现实世界的经验证据的检验：一种是看看理论是否能一般地解释观察到的现象；另一种是用数据统计检验理论的某些含义。我们下面简要地讨论这两种情况。

1. 解释能力

第一种方法是运用博弈论进行推理来解释现实中所观察到的现象，其中有许多成功的例子，最早的一个来源于国际关系领域。谢林是采用博弈论解释诸多现象的先驱，他分析了国与国之间即使没有动武的动机的情况下军备竞赛也会扩大化的问题，以及武力威胁的置信度等问题。随后，博弈论在这个领域的应用已扩展到研究一个国家在外交谈判或面临可能爆发战争的危机的时候何时以及如何表明可以置信的决心等问题。博弈论被系统地运用于经济和商业领域始于 20 世纪 70 年代中期，这种应用方兴未艾，我们这里仅讨论其中几个著名的例子。

借助博弈论，我们已经理解了行业内的厂商应该何时以及如何做出可置信的承诺以阻止新的进入者，比如，发动一场灾难性的价格战来击退行业中新的进入者。单次与重复的囚徒困境博弈让我们明白了哪种产业中竞争会很激烈，进而市场价格较低；哪种产业中能形成默许的同盟，进而市场价格和利润都较高。近年来，博弈论已经成为分析一国的政治体系和制度，以及进行跨国比较的可选工具，比如，博弈论可以解释在追求个人最终目标的过程中，如何对委员会和选举中的投票和议案进行策略性的操纵。在这本入门级的教材中，我们仅介绍一些简单的博弈论应用的例子，在本章中我们已经给出了一个例子（价格竞争），随后将介绍更多，包括分析古巴导弹危机的案例以及拍卖、投票和谈判。[9]

一些批评家对这些博弈论成功应用的例子不以为然，他们声称即使不使用博弈论，采用之前大家已知的经济学和政治学等一般原理进行分析，也能得到相同的解释。他们的观点有一定的道理，在博弈论产生之前，其中的一些分析方法就已经存在了。比如，我们曾在本章的第 1 节中讨论了两家餐馆价格竞争的均衡，这个均衡的例子在经济学中已经有一百多年的历史了；纳什均衡仅仅是对所有这类博弈的均衡概念的一般表述而已。策略投票的一些理论可以追溯到 18 世纪，“置信度”的概念早在修昔底德（Thucydides）的《伯罗奔尼撒战争史》（*Peloponnesian War*）中就有了。然而，博弈论将所有的这些应用都统一了起来，因而促进了新的发展。

在过去的30年间，涌现出了一些新的思想和应用。比如，我们现在知道了不同的拍卖方式（英式、荷式、密封投标和公开竞价）如何导致了竞标策略和拍卖者拍卖所得的差异，二次打击能力的存在如何降低了突然袭击的威胁，以及政府如何成功地操纵财政政策和货币政策以提高赢得连任的可能性。如果这些例子都可用博弈论产生之前的方法分析，那么它们可能很早以前就被解释了。

2. 统计检验

第二种检验经验证据的方法是数量和统计的方法。这种方法一般的步骤是先假设纳什均衡普遍存在，然后根据这个假设，导出表明各个变量——参与人选择的策略和结果——之间关系的方程，这些变量在具体研究的环境中都是可观察的，方程可用真实的数据进行估计。产业经济学认为，厂商之间通过设定如本章例子所述的产量和价格来展开竞争。除此之外，它们还有其他的策略选择，其中包括产品质量、投资、研发等。产量和价格的选择可以放在静态环境（某一时刻）中进行研究，而诸如投资或研发等策略竞争的博弈则是动态的。已经有许多关于这两类博弈的研究[10]，其成果都比较鼓舞人心。基于纳什均衡及其动态一般化的博弈论模型较好地拟合了许多主要产业的数据，如汽车制造业等，并且相比于原来假设完全竞争来估计供给需求曲线的分析方法，这种方法更好地解释了竞争的决定因素。

在政治上，立法过程中关于各种议题的投票都是立法委员们策略互动的结果，博弈的均衡依赖于立法委员们潜在的偏好。在美国国会中，详尽的投票记录是可以公开的信息，根据偏好和纳什均衡下投票的关系，这些信息可以用于推测立法委员的偏好。这种方法被基斯·普尔（Keith Poole）和霍华德·罗森塔尔（Howard Rosenthal）采用并取得了显著的成功。[11]他们发现美国的政治可以由两类问题的冲突来概括：一类是经济不平等；另一类则是种族歧视。

哈佛商学院教授潘卡·凯玛瓦特（Pankaj Ghemawat）提出了数量分析的混合模式，这种方法结合运用了对单个公司或产业的案例分析和对大量数据样本的统计分析。[12]他的博弈理论模型明显地提升了我们对若干关于定价、能力、创新等疑难商业决策的认识，但是他的研究所需构建的模型要考虑足够多的细节，以便公平地处理所分析的公司或产业的环境。和其他初级入门教材一样，本书回避了关于这些高级建模技术的讨论，但是我们将让你进一步接触这些方法。在第14章中，我们将分析一个源于国际政治领域的案例，用以说明这种方法。

3. 现实世界中学习的例子

最后，我们给出一个对现实世界中博弈均衡和学习过程的有趣解说，以此作为总结。人们经常参与博弈，当发生的环境不是实验室或教室的时候，博弈常常关系到他们的切身利益，使得他们有很强的动机和大量的机会来学习。史蒂芬·杰伊·古尔德（Stephen Jay Gould）找到了一个很好的例子[13]，他发现在20世纪的大部分时间里，棒球季赛所记载的最佳击球率呈下降的趋势，尤其是平均击球率在0.400或以上的球员，过去的要比现在的多。棒球迷们常常带着怀旧的情绪解释下降的原因："曾经有伟大的球员。"然而，简单一想，我们不禁要问："为什么过去没有掷球手与之抗衡，使得他们的击球率很低呢?"古尔德的解释更具系统性，他指出我们要全面地看待平均击球率，而不仅仅是最好球员的击球率。与过

去相比，现在最差击球手的击球率也有一定提高，在主联盟赛中击球率低于 0.150 的球员也比原来减少了。他将这种整体差异的缩小称为标准化或稳定化效应：

> 当棒球还是一项新兴的运动时，杀伤力很强的古怪招数层出不穷，往往令人防不胜防。威·威利·基勒（Wee Willie Keeler）可以声东击西（并在 1897 年创造了击球率 0.432 的纪录），因为接球手常常不知道他们要往哪边掷。慢慢地，球员们逐渐掌握了最佳的跑位、防守、掷球和击球的技术——变化也不可避免地减少了。最佳球员常常会遇到最为精锐的对手，要赢得胜利，常常就要挑战自己的极限，所以赛场上的伤病也更多了。

换言之，对抗策略经过一系列的调整后，这个系统就收敛到（纳什）均衡了。

古尔德整理了数十年的击球数据，发现整体的差异的确是在缩小，除了偶然的“跳跃”以外。事实上，这些“跳跃”也证实了他的理论，因为“跳跃”的出现都紧跟在均衡被外来的变化打破之后。无论是比赛规则的改变（“好球”区域扩大或缩小，掷球位置降低，或者是在赛季扩张的时候，有许多新球队或新球员加入），还是技术的改变（弹性更好的球被采用，或者铝制的球将来有可能被允许用于比赛），原先互为最优反应的体系都会偏离均衡。随着球员们的不断尝试，相互差异也暂时加剧，一些取得成功而另一些不幸失败。最后，新的均衡又达成，差异再次缩小。这一切正如我们运用学习过程和调整到纳什均衡的分析框架中所预见的一样。

我们将在第 8 章中介绍关于混合策略的证据，并在后续章节的相应位置讨论一些特定类型博弈——如囚徒困境、谈判、拍卖——的证据。到目前为止，我们所提供的实验和经验证据足以让我们对使用纳什均衡持有谨慎乐观的态度，纳什均衡可以作为首选方法或分析的出发点。整体上来说，如果你所分析的博弈是经常发生的，且参与人的数目以及博弈的规则和条件都是稳定的，那么你就可以放心地使用纳什均衡的概念。但是，如果博弈是新出现的，或者只进行过一次且参与人没有什么经验，那么这时候运用纳什均衡的概念就要更加谨慎，哪怕观察到结果不是你所计算出的均衡也不要惊讶。即使这样，你开始分析的第一步仍可以是寻找纳什均衡，然后判断这个结果是不是合理的均衡；如果不是，继续思考为何不是，通常的原因都是你对参与人的目标理解有误，而不是在他们目标正确的情况下，参与人的行动选择有误。

5.3 对纳什均衡概念的批评性讨论

除了经验证据对纳什均衡概念的批评以外，还有一些理论上对该概念的批评。本节中，我们通过举例来简要回顾一些对纳什均衡概念的理论批评和对这些批评的反驳。[14]其中，一些批评是自相矛盾的，一些则是缺乏对博弈本身的更好理解，而另一些则告诉我们纳什均衡的概念本身也存在某些不足之处，并且提出了一些扩大化或放松了的概念。这些概念具有更好的性质，我们在本小节中就介

绍一个这样的替代概念，并且提示在后续各章中将要出现的其他概念。我们相信这些介绍将使你在运用纳什均衡的概念时更加自信，同时又不失审慎。然而，一些疑难问题仍未得到解决，这也说明博弈论不是一门停滞的学科。疑难问题将给研究博弈论的年轻学者们以鼓励，因为这意味着在该领域中有新思考和新研究的广阔空间。一门所有问题都被解决了的学科是一门“死亡了”的学科。

首先，我们讨论纳什均衡概念的基本优点。本书中的大多数博弈都是非合作博弈，即各个参与人独立地选择行动。因此，可以很自然地认为，在给定其他参与人选择不变的情况下，如果某参与人的行动不是基于自己的价值体系（支付等级）最优的，那么他将会改变行动。换句话说，人们倾向于假设每个参与人的行动都是对所有其他人行动的最优反应。纳什均衡就具有这个“同时最优反应”的性质——事实上，这就是它的定义。如果某个最终结果不是纳什均衡，那么其中至少有一个参与人可以通过改变行动来获得更高的支付。

借助上述观点，杰出的博弈论大师罗杰·迈尔森（Roger Myerson）驳斥了那些凭直觉认为人们的选择有别于纳什均衡策略的观点。他反驳的方法很简单，就是将证明的担子交给纳什均衡的批评者，他说：“当被问到为什么博弈中的参与人应该采取纳什均衡策略的时候，我喜欢的回答就是问他们‘为什么不是这样呢’，并且让这些挑战纳什均衡概念的人讲清楚参与人应该如何做。如果他的回答不是纳什均衡，那么……我们可以揭示出如果参与人认为那就是对其他参与人行为的准确写照，那个结果本身就是站不住脚的。”[15]

5.3.1 在纳什均衡中对风险的处理

一些批评者认为纳什均衡的概念没有对风险给予应有的重视。在一些博弈中，人们可能认为非纳什均衡策略更安全，所以他们会选择这些策略。我们举两个这样的例子，第一个是加州伯克利大学经济学教授约翰·摩根（John Morgan）提出的，博弈表如表 5—1 所示。

表 5—1　　存在可疑纳什均衡的博弈

		列		
		A	B	C
行	A	2，2	3，1	0，2
	B	1，3	2，2	3，2
	C	2，0	2，3	2，2

逐个单元格检查，很快就发现这个博弈只有一个纳什均衡，即（A，A），其支付是（2，2）。但是，你可能会像摩根实验中的参与人一样，认为选择 C 更有吸引力，理由如下：选择 C 能保证你得到和纳什均衡时一样的支付，即 2；其次，如果你选择纳什均衡上的策略 A，只有在其他人也选择 A 的情况下你才能得到 2，那么为什么要碰运气呢？再者，如果你认为其他人可能根据这种推理选

择 C，那么你选择 A 将会犯严重的错误，此时只能得到 0，而你选择 C 可以保证得到 2。

迈尔森回答道："且慢，如果你真的相信其他人会如此认为并将选择 C，那么你应该选择 B 以求得到支付 3；如果你认为其他人会想到这点并因而选择 B，那么你对 B 的最优反应则是 A；如果你认为其他人也会想到这一点，那么你就应该选择对 A 的最优反应，即选择 A，这样又回到了纳什均衡！"正如你所看到的，批评纳什均衡和反驳这些批评本身就是一种智力上的博弈，非常引人入胜。

第二个例子则更有趣，是戴维·克雷普斯提出来的，支付矩阵如表 5—2 所示。在开始对这个博弈进行理论分析之前，请假想你自己也真正参与了这个博弈，你就是参与人 A，那么你将选择你的两个行动中的哪一个呢？

请记住你对前面问题的回答，然后我们开始分析这个博弈。首先，我们寻找占优策略，发现参与人 A 没有占优策略，而参与人 B 则有占优策略。无论 A 如何选择，B 选择"左"都能保证获得支付 10，而选择"右"只能得到 9.9（仍然是无论 A 如何选择）。因此，参与人 B 应该选择"左"。给定参与人 B 要选择"左"，参与人 A 则选择"下"更好。那么，这个博弈的唯一纯策略纳什均衡是（下，左），每个参与人在此结果上获得 10 的支付。

这里的问题是，许多（不是所有）人如果处于参与人 A 的位置上，是不会选择"下"的（你又是如何选择的呢?），尤其是那些学习博弈论多年的学生和没有接触过这门学科的人。如果 A 对 B 的支付或理性表示怀疑，那么 A 选择"上"比选择纳什均衡策略"下"更加安全。如果 A 认为支付表如表 5—2 所示，但实际上 B 的支付却是反过来的——选择"左"得到 9.9，选择"右"得到 10，情况又会怎么样呢？如果 9.9 仅是一个约数，而实际的支付是 10.1，结果又会如何呢？如果 B 有着完全不同的价值体系，或者不是一个理性人，可能为了开玩笑选择"错误"的行动，情况又会怎么样呢？很明显，完美信息和理性的假设在对策略的分析中很重要，参与人对上述假设的怀疑不仅会改变均衡从而使之有别于我们通常所做的预测，而且会对纳什均衡概念的合理性提出疑问。

表 5—2　　　　灾难性的纳什均衡？

		B	
		左	右
A	上	9，10	8，9.9
	下	10，10	−1 000，9.9

然而，这些例子的真正问题不是纳什均衡概念不适用了，而是这些例子将纳什均衡概念运用得过于简单了。在前面的例子中，如果对 B 的支付产生怀疑，那么就应该让其成为分析的一部分。如果 A 不知道 B 的支付，这个博弈就是信息不对称的，我们将在第 9 章中学习这种博弈的一般分析方法，但下面将给出这种博弈的一个相对简单的例子，我们可以轻松地计算出它的均衡。

设想 A 认为 B 选择"左"和"右"的支付与如表 5—2 所示的支付颠倒的概率是 p，那么 $(1-p)$ 就是 B 的支付如表 5—2 所示的概率。因为 A 要在不清楚

到底哪一种是真的的情况下选择行动，所以他要选择的策略是“平均上最优”的。在这个博弈中，计算很简单，因为在每种情况下B都有一个占优策略。对于A来说，唯一的问题是两种情况下B的占优策略是不同的：（在如表5—2所示的情况下）B的占优策略是“左”的概率是 $(1-p)$，（在相反的情况下）B的占优策略是“右”的概率是 p。因此，如果A选择“上”，那么他将有 $(1-p)$ 的概率遇到B选择“左”，并得到9的支付；有 p 的概率遇到B选择“右”，并得到8的支付。所以按统计或概率加权，A选择“上”的平均支付是 $9(1-p)+8p$。同样地，A选择“下”的平均支付是 $10(1-p)-1\,000p$。当下式满足时，A选择“上”会更好些：

$$9(1-p)+8p>10(1-p)-1\,000p \text{ 或 } p>1/1\,009$$

因此，即使B的支付只有很小的概率与表5—2中所示的相反，选择“上”对于A来说也是最优的。在这样的情况下，基于理性行为的分析如果正确的话，是不与直觉猜想和实验证据相矛盾的。

在前面的计算中，我们假设参与人A在面对支付的不确定性时，会分别计算不同行动所产生的统计平均支付，并且选择那个带给他最高平均支付的行动。这个隐含的假设，虽然契合我们列举这个例子的目的，但并不是没有问题的。比如，一个人如果面对两种情况，一种是各有50%的概率赢得或输掉10美元，另一种是各有50%的概率赢10 001美元和输10 000美元，那么这个假设就意味着他会选择后者，因为后者在统计平均上可以赢得50美分（$\frac{1}{2}\times 10\,001-\frac{1}{2}\times 10\,000$），而前者只能得到0（$\frac{1}{2}\times 10-\frac{1}{2}\times 10$）。但是大多数人都会认为第二种情况风险更大，因而偏好第一种情况。这个问题是容易解决的，在第7章的附录中，我们将介绍如何构造一个金钱数额上非线性的支付结构，使得决策者可以同时考虑风险和收益。

5.3.2 纳什均衡的多重性

另一种对纳什均衡概念的批评基于对许多博弈存在多个均衡这一事实的观察，因而批评者们声称纳什均衡不能足够准确地确定博弈结果并给出唯一的预测。这种观点并不意味着我们要抛弃纳什均衡，而是说明该概念本身还不够完善，还需考虑哪个均衡比其他的均衡更有理由成为博弈结果。

在第4章中，我们讨论了许多协调博弈，它们都有多个均衡。如果参与人之间存在某些共同的社会、文化、历史知识，或者博弈具有某些人为的或巧合的特征，使得人们的预期收敛，那么参与人就可以从多个均衡中选择一个作为聚点均衡。下面是一个协调博弈中纳什均衡多重性的极端例子：有100美元的总奖金，两人被要求同时且独立地写下想要得到的钱的额度，如果他们写的金额加起来不超过100美元，那么他们每个人都得到各自所写金额的钱；如果两个金额加起来超过了100美元，那么他们什么都得不到。对于任意 x，一个参与人写下 x 的同

时另一个写下（100－x）都是一个纳什均衡。因此，这个博弈有（几乎）无穷个纳什均衡。但是事实上，50∶50是一个聚点均衡。平等和公平的社会准则已经深入人心，以至于可以看成人们的一种本能，选择50的参与人认为这是一个很明显的答案。要成为一个真正的聚点，不仅要求大家都认同它，而且每个人都知道大家都认同它，以及每个人都知道……换言之，它被认同是一种共同知识。事实并不总是如此，若一个参与人是来自文明和平等社会的女子，则她认为50∶50是一种显而易见的分配方案；而另一个是来自父权制度社会的男子，则他认为无论如何分，男人得到的总要是女人的三倍。这样一来，他们都选择了自己所认同的行为，最后他们什么都得不到，因为他们各自认同的分配方案不是双方的共同知识。

聚点均衡的存在常常是一种巧合，而人为的创造聚点则是一门艺术，它需要对博弈本身的历史和文化背景有足够的认识，而不仅仅是数学描述。这对于许多博弈论学者来说都是一件头疼的事情，因为他们都更希望结果只依赖于对博弈的抽象刻画——参与人及其策略都可以用数来表示，且没有其他什么外部关联。然而，我们的看法则不是这样的，我们认为历史和文化背景对于博弈来说是和数学描述同等重要的，并且如果这些背景知识帮助我们从多个纳什均衡中挑选出一个唯一的结果，那么很难再说它不重要了。

在第6章中，我们将看到序贯行动博弈可能存在多个纳什均衡，我们通过引入“置信度”的条件，筛选出一个特定的均衡，即第3章所述的逆向归纳均衡。在信息不对称或更为复杂的博弈中，我们通过引入被称为“精炼”（refinements）的限制来识别和剔除一些不合理的纳什均衡。在第9章中，我们将介绍这种精炼过程，它可以导出被称为“贝叶斯精炼纳什均衡”的结果。就某种特定类型的博弈而言，每次精炼的目的都是具体的，它约定了当参与人观察到别人采取或没采取行动时如何修正自己的信息。在具体的情况中，这样的约定常常都是非常合理的。在许多博弈中，通过这样的方法都不难排除大部分的纳什均衡，从而使得我们的预测更加明确。

批评纳什均衡有太多均衡，反过来就是说一些博弈根本就不存在均衡。在4.8节中，就有一个这样的例子，但通过将策略的概念扩展到随机混合形式，重新又找到了纳什均衡。在第7章和第8章中，我们将介绍和解释混合策略纳什均衡。在博弈论更高深的研究中，一些难懂的例子甚至连混合策略纳什均衡都不存在，但是这些复杂的情况与本书所介绍的方法和应用无关，我们不准备加以阐述。

5.3.3 纳什均衡对理性的要求

纳什均衡被认为是一个由参与人策略选择和参与人对其他参与人行动的信念所构成的体系。在均衡上，给定对其他参与人行动的信念，每个参与人的选择都带来了最高的支付，并且每个参与人的信念是正确的，即，参与人的实际选择与大家对他行动所持的信念是一致的。这看起来是对个人理性的相互一致性条件的自然表述。如果“所有参与人都是理性的”是大家的共同知识，那么怎么会有人

认为其他参与人的选择不是对自己行动的理性反应呢？

为解决这个问题，我们首先考虑如表 5—3 所示的 3×3 博弈。逐个单元格检查，很快发现它只有一个纳什均衡，即（R2，C2），其支付是（3，3）。在均衡上，行选择 R2，因为他相信列会选择 C2，为什么他会这么认为呢？因为他相信列是理性的，同时他也相信列认为他会选择 R2。如果列认为行会选择 R1 或 R3 的话，C2 就不是列的最优选择。因此，在形成信念和反应的理性过程中，信念必须要是正确的。

表 5—3　用信念和反应的链条来说明选择合理

		列		
		C1	C2	C3
行	R1	0，7	2，5	7，0
	R2	5，2	3，3	5，2
	R3	7，0	2，5	0，7

上述观点的问题在于对信念的考虑只进行了一轮，如果我们继续下去，那么将导出其他选择的组合也是合理的。比如，我们可以证明行选择 R1 是合理的。我们知道如果行认为列会选择 C3，那么他的最优选择就是 R1。为什么行会这么认为呢？因为他认为列相信他会选择 R3。行认为这个信念是合理的，因为他认为列会相信行相信列由于相信行会选择 R1 而选择 C1，依此类推……这是一个信念的链条，每一环都是完全理性的。

因此，单纯依靠理性并不能证明纳什均衡是合理的。有更为复杂的观点，可以解释某种特定形式的纳什均衡的合理性，其中参与人是根据大众可观测的随机机制来选择自己策略的。然而，我们把这个知识留给更高级的教材去讲述。在下一节中，我们将提出一个较简单的概念，它将告诉我们单纯依靠参与人理性这一共同知识，逻辑上意味着什么。

5.4　可理性化

仅仅根据理性，博弈中哪些策略选择是合理的呢？如表 5—3 所示的支付表中的每对策略组合，我们都可借助曾在 5.3.3 节用过的相同逻辑证明它们都是合理的。换句话说，就是这九对组合在逻辑上都可以说得过去。因此，仅仅依靠理性并不足以让我们确定和预测结果。所有的博弈都是这样吗？不是的。比如说如果某个策略是劣策略，理性就可以将其剔除掉。如果参与人认识到其他人是理性的，是不会选择劣策略的，那么依据理性这一共同知识，劣策略就可被重复剔除。运用理性的概念，仅能做到这点吗？不。根据一个比在纯策略中占优稍强一点的性质，还能剔除其他的一些策略。这个性质可以识辨出**绝非最优反应**（nev-

er a best respond）的策略，剔除这些策略，剩下的策略就叫做**理性化的**（rationalizable）策略。这个剔除过程本身被称为**可理性化**（rationalizability）。

5.4.1 可理性化概念的运用

在表 5—3 的基础上，为每个参与人各加一个策略，便得到一个新的博弈，如表 5—4 所示。[16]刚才我们已经指出，根据对他人信念的信念链条，从每个参与人的前三个策略中任选一个所组成的 9 对策略组合都是合理的，在这个扩大了的支付矩阵中仍是如此。但是，能否通过这样的方法来说明 R4 和 C4 是合理的呢？

表 5—4　　　　理性化的策略

		列			
		C1	C2	C3	C4
行	R1	0，7	2，5	7，0	0，1
	R2	5，2	3，3	5，2	0，1
	R3	7，0	2，5	0，7	0，1
	R4	0，0	0，−2	0，0	10，−1

行相信列会选择 C4 吗？这个信念是否合理，依赖于列关于行的选择的信念。那么，要使得 C4 成为列的最优选择，列关于行的选择的信念应该是什么呢？答案是这样的信念不存在。如果列相信行会选择 R1，那么列的最优选择是 C1；如果列相信行会选择 R2，那么列的最优选择是 C2；如果列相信行会选择 R3，那么列的最优选择是 C3；如果列相信行会选择 R4，那么列的最优选择是 C1 或 C3。因此，C4 不可能成为列对行的行动的最优反应。[17]这就意味着如果行知道列是理性的，行就明白无论自己如何选择，都不能使列具有选择 C4 的信念，所以行相信列是永远不会选择 C4 的。

注意，虽然 C4 是绝非最优反应，但它并不被 C1、C2、C3 所占优。C4 对于列来说，在行选择 R3 的时候优于 C1，在行选择 R4 的时候优于 C2，在行选择 R1 的时候优于 C3。如果某策略是劣策略，它也不可能是最优反应。因此，"绝非最优反应"是一个比"占优"更一般的概念，即使在不能剔除劣策略的时候，也可能剔除绝非最优反应策略，所以剔除绝非最优反应策略可以在剔除劣策略的基础上进一步减少可能的结果。[18]

绝非最优反应策略的剔除也可以相继进行。如果行和列都是理性的，那么行不相信列会选择 C4，并且列也会推知这一点，而 R4 只有在列选择 C4 的时候才是行的最优反应，所以列也不相信行会选择 R4。因此，R4 和 C4 都不属于理性化策略的集合，"可理性化"让我们将可能的结果空间缩小了。

如果博弈存在一个纳什均衡，那么其均衡是理性化的，并且经得起一整套信

念体系的检验，正如我们在 5.3.3 节所见的一样。但是，更为一般地，即使博弈没有纳什均衡，它也可能有理性化的结果。在表 5—3 或表 5—4 中，行仅保留策略 R1 和 R3，列仅保留策略 C1 和 C3，考虑这样一个 2×2 的博弈，显然这个博弈没有纯策略纳什均衡，但是所有的这四个结果都是理性化的，因为这些策略都经得住信念链条的检验。

因此，对于那些没有纳什均衡的博弈而言，“可理性化”提供了一种可能的解法。更为重要的是，这个概念告诉我们，如果仅仅依据理性，我们可以将博弈中可能的结果缩小到什么范围。

5.4.2 可理性化导出纳什均衡

在一些博弈中，重复剔除绝非最优反应策略，可以将结果的可能空间缩小到只剩下纳什均衡。这一点是非常有用的，因为在这些博弈中，我们可以说纳什均衡完全是从参与人对其他人信念的理性思考中得来的。能用这种方法求解的博弈在经济学中也是非常重要的。如果厂商们选择各自生产的产量，并且它们投入市场的总的产量将决定价格，那么它们之间的竞争就属于这种类型的博弈。

我们下面介绍一个这种类型的博弈。在一个海滨小镇上，有两条渔船，它们每天晚上出去捕鱼，第二天早上回来，并将它们前一天晚上打到的鱼放到市场上去销售。这个博弈发生的时候还没有现代的电冰箱，所以打到的鱼都要在当天被卖出去和消费掉。小镇附近的海域盛产鱼类，所以渔船的主人可以决定每天晚上打多少鱼，但他们知道如果投入市场的鱼太多的话，过分的供给将意味着低的市场价格和低的利润。

我们假设，如果一条船打回 X 桶鱼投入市场，同时另一条船打回 Y 桶鱼投入市场，那么市场价格 P（以每桶若干达克特来计量）将是 $P=60-(X+Y)$。我们同时假设两条船及其船员的捕鱼效率是不同的，第一条船的捕鱼成本是每桶 30 达克特，第二条船是每桶 36 达克特。

下面我们写出这两个船主的利润 U 和 V 分别与他们的策略 X 和 Y 的关系：

$$U=[(60-X-Y)-30]X=(30-Y)X-X^2$$
$$V=[(60-X-Y)-36]Y=(24-X)Y-Y^2$$

依据这些支付的表达式，我们可以求解最优反应曲线，并且找出纳什均衡。

第一条船的最优反应 X 应该在给定另一条船的 Y 不变的情况下使得 U 最大化。运用微积分，我们可以保持 Y 不变，同时对 U 求 X 的微分，令这个微分式等于 0，我们得到

$$(30-Y)-2X=0 \text{ 即 } X=15-Y/2$$

不用微分的方法，我们将 U 的表达式化成如下平方的形式：

$$U=[(30-Y)/2]^2-[(30-Y)/2-X]^2$$

我们观察到 X 仅出现在右边的第二项平方式中。因为此项前面带一个负号，所

以要使 U 尽量大，就要令此项尽量小，即等于 0，这样得到 $(30-Y)/2-X=0$ 或 $X=15-Y/2$。

同样地，第二条船的最优反应方程是保持 X 不变的情况下找到使得 V 最大化的 Y 值，即

$$(24-X)-2Y=0 \text{ 亦即 } Y=12-X/2$$

联立求解上述两个关于 X 和 Y 的最优反应方程，很容易就得出纳什均衡，所以我们这里仅给出结果[19]：产量分别是 $X=12$ 和 $Y=6$，价格是 $P=42$，利润分别是 $U=144$ 和 $V=36$。

图 5—3 展示了两个船主的最优反应曲线（对应的方程用 BR1 和 BR2 标注），这两条曲线的交点就是纳什均衡（标注 N 及对应的支付）。同时，图 5—3 也揭示了参与人关于其他参与人选择的信念是如何随着绝非最优反应策略的重复剔除而缩小的。

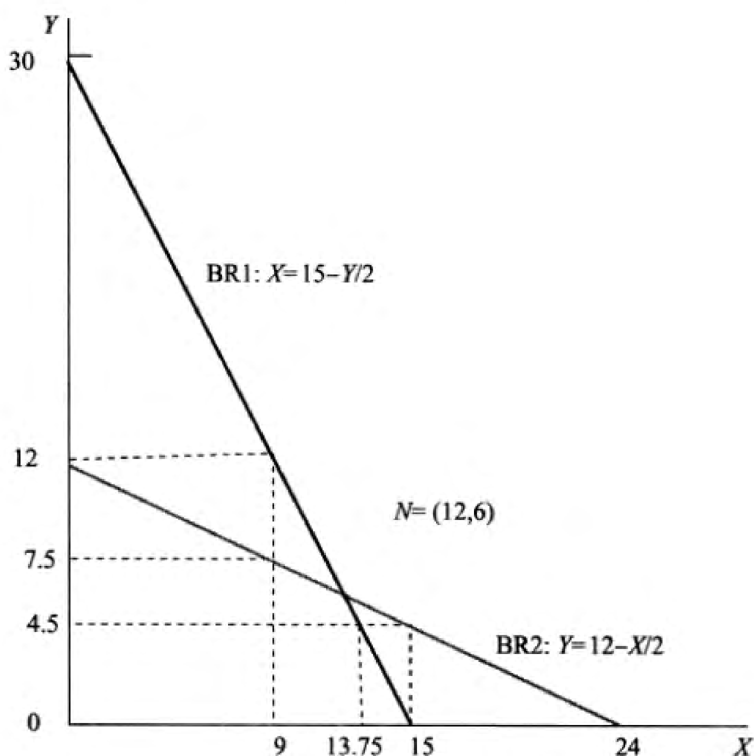


图 5—3　理性化导出纳什均衡

依据理性，第一个船主认为第二个船主选择的 Y 值是多少呢？这依赖于第二个船主认为第一个船主要捕多少鱼。但是无论怎么样，第二个船主的最优反应区域是在 0 到 12 之间，所以第一个船主认为第二个船主的选择不会超出这个区域，所有小于 0 的 Y 值（很明显）和所有大于 12 的 Y 值（不太明显）都被排除了。同样地，第二个船主也会理性地认为第一个船主的产量不会低于 0 或大于 15。

现在我们进行第二轮剔除。如果第一个船主认为第二个船主对 Y 的选择限于 0 和 12 之间，那么他自己对 X 的选择就缩小到对这些 Y 值的最优反应的区间。对 $Y=0$ 的最优反应是 $X=15$，对 $Y=12$ 的最优反应是 $X=15-12/2=9$。因为

BR1 是向下倾斜的，经过这轮下来，X 的全部范围缩小为 9 到 15 之间。同样地，第二个船主 Y 值的选择也限于对 X 取 0 到 15 之间的各值的最优反应上，即，$Y=12$ 和 $Y=12-15/2=4.5$ 之间的各值。这些限定的范围，如图 5—3 的坐标轴所示。

第三轮理性化还能继续将区域缩小，因为 X 的取值必须大于 9，同时 BR2 具有负的斜率，所以 Y 最大的取值是对 9 的最优反应，即，$Y=12-9/2=7.5$。经过第二轮分析得出，Y 值至少大于 4.5，所以 Y 值现在被限定在 4.5 和 7.5 之间。同样地，由于 Y 值至少是 4.5，X 值则最大是 $15-4.5/2=12.75$。第二轮分析得出 X 值大于 9，所以现在 X 值介于 9 到 12.75 之间。

你可以继续进行这样的理性化过程，但是结论已经很明显了，两个值域的相继缩小过程是收敛于纳什均衡 $X=12$ 和 $Y=6$ 的。因此，纳什均衡是唯一一个不被重复剔除绝非最优反应策略所剔除的。[20] 如果博弈的最优反应曲线是向下倾斜的，并且在交点处只有一个纳什均衡，那么这个结论都是适用的。[21]

我们还要谨慎地区分上述观点与过去讨论过的相继最优反应链条。相继最优反应的逻辑推理如下：从某个参与人的任意一个策略开始，比如 $X=18$，然后另一个参与人对此的最优反应是$Y=12-18/2=3$；X 对 $Y=3$ 的最优反应又是 $X=15-3/2=13.5$；反过来，Y 对 $X=13.5$ 的最优反应是 $12-13.5/2=5.25$；然后再反过来，X 对 Y 的最优反应是 $X=15-5.25/2=12.375$，依此类推。

过去我们所讨论的最优反应链也收敛于纳什均衡，但这种方法是有问题的。因为博弈只进行一次且行动是同时的，所以参与人不可能等其他人选择了才做出反应，然后再让其他人做出反应，依此类推。如果允许这种动态性存在，那么参与人为什么不在预测其他人将来反应的基础上从一开始就选择不一样的行动呢？

然而，可理性化的观点是不一样的，它明显地体现了博弈的一次性和行动的同时性等因素，所有对最优反应链的思考都是先期进行的，并且后续的各轮推理和反应都是存在于观念之中的。参与人不是对真实的行动做出反应，而是仅仅推算他们将做的选择，所以这种动态性只是存在于参与人的思想中。

5.5 总　结

在同时行动博弈中，如果参与人在一个连续的区间内选择行动，我们就可以运用最优反应分析来求得数学形式的最优反应规则，将其联立求解进而得到纳什均衡策略。最优反应规则也可以用曲线来表示，两条曲线的交点代表的就是纳什均衡。厂商在一个较广的值域中选择价格或产量，政党决定其竞选宣传的支出，这些例子都是连续型策略的博弈。

关于纳什均衡的实验结果显示，共同的文化背景对存在多个均衡的博弈中的协同是非常重要的，重复进行一些博弈表明参与人能在经验中学习并逐渐选择趋于纳什均衡的行动。另外，均衡预测的准确性还取决于实验设计者的假设是否与参与人的真实偏好相匹配。博弈论在真实世界的应用尤其有助于经济学家和政治

学家理解消费者、厂商、选民、议员和政府的行为。

对纳什均衡概念的理论批评认为纳什均衡没有充分考虑风险，并且由于许多博弈存在多个均衡，纳什均衡的应用是有限的，此外，仅依靠理性不能导出纳什均衡。在许多情况下，对博弈及其支付进行更好的描述，或者对纳什均衡进行精炼，都可以使预测更好或者减少可能的均衡。“可理性化”是指剔除那些绝非最优反应策略，进而得到一系列理性化的结果。当博弈有一个纳什均衡时，这个结果是理性化的；但是，理性化仍然可以预测那些没有纳什均衡的博弈的结果。

关键术语

最优反应曲线（best-response curves）
最优反应规则（best-response rule）
连续型策略（continuous strategy）
绝非最优反应（never a best response）
可理性化（rationalizability）
理性化的（rationalizable）
精炼（refinement）

习　　题

1. 在 5.1.2 节的政党竞选宣传博弈中，政党 L 将宣传费用定为 x（百万美元），同时政党 R 定为 y（百万美元），我们还得出了这个博弈中政党 R 的最优反应规则是 $y=\sqrt{x}-x$，政党 L 的最优反应规则是 $x=\sqrt{y}-y$。运用这些最优反应规则来验证纳什均衡的宣传费用是否是 $x=y=1/4$ 或 250 000 美元。
2. 图 5—1 所示的餐馆博弈定义了泽维尔（Q_x）的需求函数和伊冯娜（Q_y）的需求函数，分别是 $Q_x=44-2P_x+P_y$ 和 $Q_y=44-2P_y+P_x$。另外，每家餐馆的利润还取决于它们服务每个顾客的成本。现在假设伊冯娜通过重新安排工作和裁掉一些员工将服务每个顾客的成本降到了 6 美元；而泽维尔服务每个顾客的成本仍然保持为 8 美元不变。
   (a) 在新的成本条件下，重新计算最优反应规则和纳什均衡时两家餐馆的价格。
   (b) 画出两家餐馆的最优反应曲线，说明你画的图和图 5—1 的区别，尤其请指出哪一条线移动了以及移动了多少，你能解释图中的这种变化吗？
3. 雅皮镇（Yuppietown）有两家食品店，分别是卖面包的法式饼店（La Boulangerie）和卖奶酪的风味餐店（La Fromagerie）。生产一块面包的成本是 1 美元，生产一磅奶酪的成本是 2 美元。如果 La Boulangerie 销售每块面包的价格是 P_1 美元，同时 La Fromagerie 销售每磅奶酪的价格是 P_2 美元，那么它们各自的销量 Q_1（千块面包）和 Q_2（千磅奶酪）分别满足下面的表达式：

   $Q_1=10-P_1-0.5P_2$，$Q_2=12-0.5P_1-P_2$

(a) 找出这两家商店的最优反应规则，画出最优反应曲线，并求出这个博弈的纳什均衡价格。

(b) 假设这两家店合谋并联合定价以最大化它们总的利润，求出实现两家店共同利润最大化的价格。

(c) 凭直觉解释纳什均衡价格和实现共同利润最大化的价格之间的差异，为什么共同利润最大化的价格不是纳什均衡呢?

(d) 在这个问题中，面包和奶酪互为互补品，它们常常一起被消费，所以它们中的一种降价可以促进另一种的销量，而 5.1.1 节中的两个餐馆的产品互为替代品。你在这里得出的最优反应规则、纳什均衡价格和共同利润最大化价格与前面餐馆定价博弈例子中所得到的结论是不同的，互补品和替代品之间的差异是如何说明这种不同的呢?

4. 在里约热内卢（Rio de Janeiro）的海滩上，有两辆推车销售椰奶（在椰子里），它们之间相隔 1 英里，并且分别位于 0 英里处和 1 英里处（在这段海滩上，只有这两辆卖椰奶的推车）。这两辆推车——推车 0 和推车 1——卖出每个椰奶的价格分别为 p_0 和 p_1，买的人是海滩上的游客，并且均匀地分布在这段海滩 0 英里处和 1 英里处之间。每个游客在一天中都要买一个椰奶，除了价格，他们还得支付运送费用 0.5 乘以 d^2，其中 d 是从他的沙滩毯到卖椰奶的推车之间的距离（单位是英里）。这样一来，推车 0 的顾客是处于 0 英里处和 x 英里处之间的游客，而推车 1 的顾客是处于 x 英里处和 1 英里处之间的游客，其中 x 是去这两个推车买椰奶的费用无差异的游客所处的位置。x 满足以下的表达式：

$$p_0+0.5x^2=p_1+0.5(1-x)^2$$

两辆推车将设定销售椰奶的价格，以最大化它们的利润 B。利润由收益（销售椰奶的价格乘以卖出的数量）和成本（每个椰奶的成本 0.25 美元乘以卖出的数量）共同决定。

(a) 写出光顾每辆推车的顾客人数的表达式（记住推车 0 的顾客来自于 0 英里处和 x 英里处之间，或者是 x；而推车 1 的顾客来自于 x 英里处和 1 英里处之间，或者是 $1-x$）。

(b) 写出两辆推车的利润函数，并且找出它们各自定价的最优反应规则。

(c) 画出最优反应曲线，然后计算（或在你的图上标出）纳什均衡时海滩上椰奶的价格。

5. 表 5—1 所示的博弈有一个唯一的纯策略纳什均衡，但是所有的九个博弈结果都是理性化的。证明这个论断，并且对每个结果说明你的推理。

6. 表 5—2 所示的博弈有一个唯一的纯策略纳什均衡，找出这个纳什均衡，并且说明它也是博弈中唯一的一个理性化的结果。

7. 5.4.2 节讲述了一个发生在海滨小镇上的捕鱼博弈，推导出两个船主的最优反应规则后，可理性化可以说明这个纳什均衡是合理的。在前文中，我们通过三轮理性化过程剔除了一些绝非最优反应策略。到了第三轮，我们知道 X（第一条船捕的鱼的桶数）至少为 9，以及 Y（第二条船捕的鱼的桶数）至少

为 4.5，第三轮剔除将 X 的取值缩小到 9 到 12.75 之间，同时将 Y 的取值缩小到 4.5 到 7.5 之间。请在此基础上再继续一轮（即第四轮）可理性化的剔除，求出经过这一轮后的缩小了的 X 和 Y 的区域。

8. 在最优反应曲线是向上倾斜的博弈中，如果绝非最优反应策略的剔除从最小的可能值开始，纳什均衡可以通过可理性化过程得到。考虑如图 5—1 所示的泽维尔的达帕斯西班牙餐馆和伊冯娜的炭火西餐馆之间的价格竞争博弈，使用该图及其背后的最优反应规则来证明这个博弈的纳什均衡是可理性化的。从两家餐馆的最小的可能价格开始，（至少）描述两轮理性化价格区间缩小并趋于纳什均衡的过程。

9. （选做，需要运用微积分）记得在 5.1.2 节政党 L 和政党 R 的竞选宣传博弈例子中，若 L 花费 x 百万美元宣传，同时 R 花费 y 百万美元，那么 L 得到的选票份额为 $x/(x+y)$，R 得到的选票份额为 $y/(x+y)$。在这个模型中，我们曾提到政党间可能有两种类型的不对称性存在，一个政党——比如 R——可能以较低的成本宣传，或者 R 的宣传支出比 L 在赢得选票上更有效率。考虑到这两种可能，我们将两个政党的支付函数写成如下形式：

$$V_L=\frac{x}{x+ky}-x \text{ 以及 } V_R=\frac{ky}{x+ky}-cy$$

这两个支付函数的形式说明：当 k 比较大的时候，R 在宣传上有效率优势；当 c 比较小的时候，R 在宣传的成本上有优势。

(a) 运用支付函数导出 R（选择 y）和 L（选择 x）的最优反应方程。

(b) 运用你的计算器或电脑，画出 $k=1$ 和 $c=1$ 时各自的最优反应方程，并比较这个图与 $k=1$ 和 $c=0.8$ 时的差异，宣传成本的优势有什么效应呢？

(c) 比较（b）中得到的当 $k=1$ 和 $c=1$ 时的图与 $k=2$ 和 $c=1$ 时的图的差异，宣传效率上的优势有什么效应呢？

(d) 联立你在（a）中得到的最优反应方程，解出 x 和 y，验证纳什均衡时的竞选宣传支出是

$$x=\frac{ck}{(c+k)^2} \text{ 和 } y=\frac{k}{(c+k)^2}$$

(e) 令纳什均衡时支出方程中的 $k=1$，并且考察两个均衡支出是如何随 c 的变化而变化的。然后再令 $c-1$，考察这两个均衡支出是如何随着 k 的变化而变化的。你的答案是否支持你在本习题（b）和（c）中所得到的结论呢？

【注释】

[1] 事实上，竞争会持续一段时期，所以每个参与人都可以观察到对方过去的选择。博弈的这种重复进行将带来新的思考，我们将在第 11 章中介绍。

[2] 懂得一点经济学知识的读者知道，联系价格和数量的方程是两种产品 X 和 Y 的需求方程。每种产品的需求数量随着自身价格的上升而下降（此时，需求曲线是向下倾斜的），随着另一种产品价格的上升而上升（此时，这两种产品是替代品）。

[3] 如果没有这种对称性，两个最优反应方程是不同的，但鉴于我们的假定，仍然是线性的。所以，在不对称情形下求解也不是那么困难。

[4] 公司知道可以逃脱合谋的处罚时，会试图进行明显的合谋，在 Kurt Eichenwald 写的 *The Informant*（New York：Broadway Books，2000）中就有这样一个有趣并富有意义的例子。

[5] Douglas D. Davis and Charles A. Holt, *Experimental Economics*, Princeton: Princeton University Press, 1993.

[6] Thomas Schelling, *The Strategy of Conflict*, New York: Oxford University Press, 1960, pp. 54 - 58; David Kreps, *A Course in Microeconomic Theory*, Princeton University Press, 1990, pp. 302 - 303, 414 - 415.

[7] 著名的博弈理论家 Jorgen Weibull 对此观点有详细的介绍，请见"Testing Game Theory", Working Paper EFI 382, Stockholm School of Economics, 2000。

[8] 如果你记入你自己的选择，这个计算就会更完备。假如有 N 个参与人，在"最差的情形"下，其他的（$N-1$）个参与人都选择 100，只有你选择 x，平均数就是$[x+(N-1)100]/N$，那么你的最优选择就是这个数的一半，即 $x=[x+(N-1)100]/(2N)$或 $x=100(N-1)-(2N-1)$。如果 $N=10$，那么$x=900/19=47$（约数）。所以任何高于 47 的选择都劣于 47，依此类推。

[9] 读者若想了解更多应用，我们这里提供一些建议的文献。Thomas Schelling 写的 *Strategy of Conflict*（New York: Oxford University Press, 1960）和 *Arms and Influence*（New Haven: Yale University Press, 1966）仍然是所有学习博弈论的学生必读的书目。采用博弈论解释产业组织理论的经典教材是 Jean Tirole 著的 *Industrial Organization*（Cambridge: MIT Press, 1988）。在政治学中，早期的经典文献是 William Riker 著的 *Liberalism Against Populism*（San Francisco: W. H. Freeman, 1982），近期文献请参考 Robert J. Aumann 和 Sergiu Hart 编撰的 *The Handbook of Game Theory*（Amsterdam: North-Holland, 1992, 1994, 2002）中的几篇文章，特别是第二卷中 Barry O'Neill 的文章 *Game Theory Models of Peace and War* 和第三卷中 Jeffrey Banks 的文章 *Strategic Aspects of Political Systems*。

[10] 关于价格和产量竞争的静态研究的文献综述，参见 Timothy F. Bresnahan, "Empirical Studies of Industries with Market Power", in *Handbook of Industrial Organization*, ed. Richard L. Schmalanesee and Robert D. Willig Vol. 2, Amsterdam: North-Holland, 1989。Ariel Pakes 在 *A Framework for Applied Dynamic Analysis in Industrial Organization*（National Bureau of Economic Research Working Paper No. 8024, Cambridge, MA: NBER, December 2000）中介绍了动态研究的一般方法。

[11] 参见 Keith Poole and Howard Rosenthal, "Pattern of Congressional Voting", *American Journal of Political Science*, Vol. 35, no. 1, February 1991, pp. 228 - 278; *Congress: A Political-Economic History of Roll Call Voting*, New York: Oxford University Press, 1996。

[12] Pankaj Ghemawat, *Games Businesses Play: Cases and Models*, Cambridge: MIT Press, 1997.

[13] "Losing the Edge," in *The Flamingo's Smile*, New York: Norton, 1985, pp. 215 - 229.

[14] David M. Kreps 在 *Game Theory and Economic Modelling*（Oxford: Clarendon Press, 1990）中对此有非常不错的讨论。

[15] Roger Myerson, *game Theory*, Cambridge: Harvard University Press, 1991, p. 106.

[16] 这个例子源于提出了"可理性化"这个概念的原始论文，参见 Douglas Bernheim, "Rationalizable Strategic Behavior," *Econometrica*, vol. 52, no. 4, July 1984, pp. 1007 - 1028。还可以参考 Andreu Mas-Colell, Michael Whinston, and Jerry Green, *Microeconomic Theory*, New York: Oxford University, 1995, pp. 242 - 245。

[17] 注意在这个例子中，对于行来说其最优选择是严格优于 C4 的，所以 C4 下的支付甚至都不可能与最优反应下的支付相等。我们也可以像区分强的和弱的占优一样，来区分强的和弱的绝非最优反应。这里，我们指的是强的绝非最优反应。

[18] 第 8 章，我们将看到在二人博弈中绝非最优反应策略劣于其他某些策略的随机组合，所以在允许混合策略的二人博弈中，这两种剔除的概念是等价的。

[19] 虽然得到这样的结果是偶然的，但是这些解的一些性质值得指出来。由于成本不同，产量也是不同的；生产效率越高（成本越低）的船销售的鱼也越多。成本和产量的差异一起意味着最终利润的更大差异，第一条船较第二条船仅有 20%的成本优势，但它的利润却是第二条船的 4 倍。

[20] 这个例子还可以用相继剔除占优策略来解，但是占优的证明较困难且需要更多的微积分知识，而绝非最优反应性质在图 5—3 上是很明显的，所以我们采用这种方法更简单。

[21] 对于向上倾斜的最优反应曲线而言，也有类似的结论。比如在图 5—1 的餐馆定价博弈中，可以从一个低的价格开始缩小最优反应的区间。只有存在明显的起始点时，才能限定最高的边界。这个起始点可能是由于外在的某些原因不能超过一个很高的价格，比如说人们没有支付高于某个价格的能力。

第 6 章 同时行动与序贯行动并存的博弈

第 3 章讨论了纯粹序贯行动博弈的本质，第 4 章和第 5 章则讨论了纯粹同时行动博弈的本质。同时，我们讨论了针对纯粹博弈类型的相关概念与分析技术——针对序贯行动博弈的博弈树与逆推法，针对同时行动博弈的支付表与纳什均衡。然而，事实上许多博弈中同时包括这两种互动。尽管我们只用博弈树（扩展式）说明序贯行动博弈，或只用支付表（策略式）来说明同时行动博弈，我们仍然有可能用这两种方式中任何一种来说明所有类型的博弈。

在本章里，我们验证了许多种可能性。我们首先展示了如何将博弈树与支付表相结合解决同时行动与序贯行动并存的博弈，以及如何将逆推法与纳什均衡相结合解决类似情况。然后，我们考察了在一个特殊博弈中改变互动性质的影响。特别地，我们关注了把博弈规则改变为同时行动中融入序贯行动或在序贯行动中融入同时行动，以及在序贯行动博弈中改变行动顺序所造成的影响。这些主题使得我们有机会将在序贯行动博弈中通过逆推法得到的均衡与在同时行动博弈中通过纳什均衡概念得到的均衡进行比较。通过这些比较，我们将纳什均衡的概念拓展到了序贯行动博弈中，从而我们证明了逆推均衡是纳什均衡的一个特例，通常称为纳什精炼均衡。

6.1 同时行动与序贯行动并存的博弈

正如前面提到过多次的，现实生活中碰到的真实的博弈大多是由若干部分构成，每一部分都可能伴随着同时行动博弈或序贯行动博弈，因此，要了解整个博弈过程就需要对两者都非常熟悉。最明显的关于同时行动与序贯行动并存的博弈的例子是那些两个（多个）参与人之间进行的多阶段博弈。例如，在一周内，你有可能同你的室友进行一系列的同时行动博弈；你在前一阶段的行动对于你决定下一阶段的行动是非常重要的，而对于你的室友也一样。同样，许多体育竞技项目、同一行业中竞争企业间的互动以及政治斗争都是由一系列同时行动博弈行为序贯连接起来的。把第 3 章所介绍的工具（博弈树与逆推法）与第 4 章、第 5 章所介绍的工具（支付表与纳什均衡）相结合可以用来分析这种博弈。[1]唯一的不同在于随着互动行为数量的增加，分析过程会变得愈益复杂。

6.1.1 两阶段博弈与子博弈

我们用以阐释的主要例子是两家想要成为电信巨人的公司——串音（CrossTalk）公司与全球对话（GlobalDialog）公司——之间的博弈。它们必须同时决定是否投资 100 亿美元采购光纤网络产品。如果两者都决定不投资，博弈结束。如果一家投资，另一家不投资，那么投资的公司具有该项电信服务的价格决定权。它可以选择制定一个较高的价格，从而吸引 6 000 万顾客，每名顾客能够带来 400 美元收益；它也可以选择制定一个较低的价格，从而吸引 8 000 万顾客，每名顾客能够带来 200 美元收益。如果两家公司都选择了投资并进入了市场，它们的定价策略就变成了一个两阶段同时行动博弈。每一家都可以选择低价或者高价。如果它们都选择高价，它们将平分市场，分别获得 3 000 万顾客，每名顾客能够带来 400 美元收益。如果它们都选择低价，它们也将平分市场，分别获得4 000万顾客，每名顾客能够带来 200 美元收益。如果一家公司选择了高价而另一家选择了低价，那么选择低价的公司将获得全部 8 000 万顾客，同样每名顾客能够带来 200 美元收益，而选择高价的公司将一无所获。

串音和全球对话两家公司的互动行为就形成了一个两阶段博弈。在第一阶段同时行动的决策中包含四种行动组合，其中一种情况导致博弈结束，另外两种情况则由一个参与人进入第二阶段进行定价决策，第四种情况则导致在第二阶段进行同时行动（定价）博弈。整个博弈过程如图 6—1 所示。

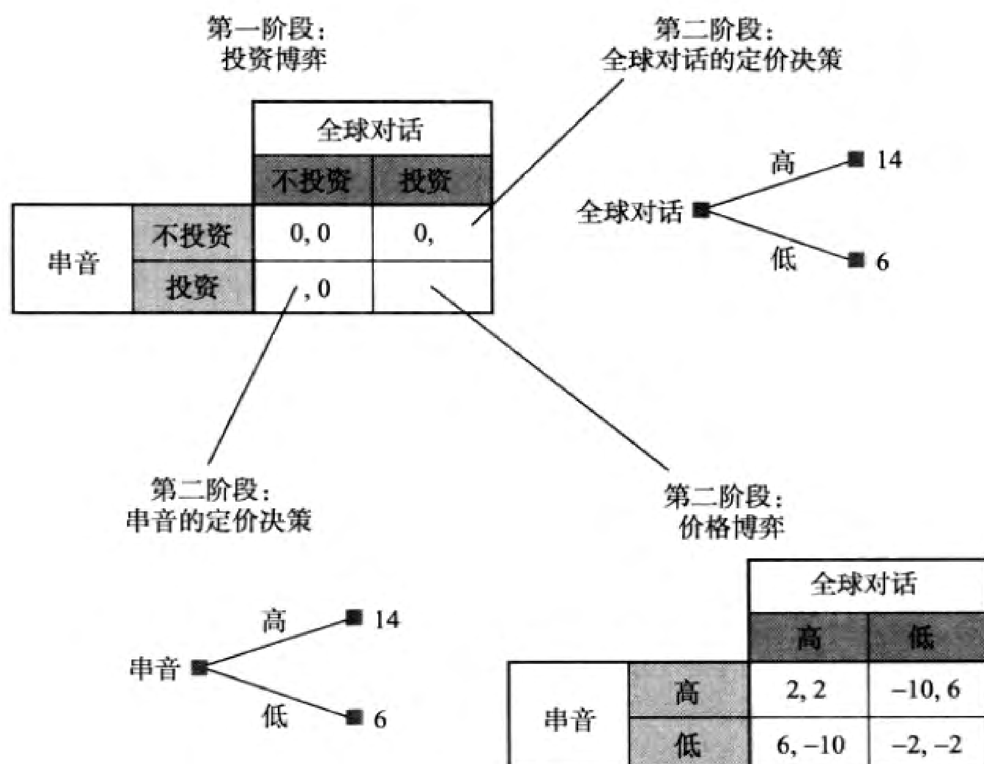


图 6—1　同时行动与序贯行动并存的两阶段博弈

图 6—1 展示的博弈树是描述整场博弈的，它比第 3 章中的博弈树更加复杂。你可以把它理解为一个具有多个层次的精心制作的“树房”。这些层次被表示在同一个二维空间表格的不同部分，就仿佛在一架盘旋在树顶上空的直升机上向下俯视一样。

图 6—1 左上方的支付表表示了第一阶段的博弈。你可以把它看成“树房”的底层。它拥有四个“房间”。西北角的“房间”表示了在第一阶段双方都不投资的情况。如果公司的决策使得博弈进入了这个房间，就不会有下一步的选择；因此我们可以把它当做第 3 章中的博弈树里的终点结，并把相关支付表示在表格里，即两家公司都获得收益 0；然而，两家公司的其他行动组合会导致博弈进入具有进一步选择的其他“房间”。因此，不能够在这些“房间”里表示出相应的支付情况。我们只好用树枝指引我们进入第二层。在东北和西南的“房间”里只标示了未投资公司所获得的支付，而那些树枝则把我们从底层的“房间”里带到了第二阶段单一公司的定价决策里。东南角的“房间”则通向“树房”中拥有多个“房间”的第二层，它表示两家公司在第一阶段都选择投资时在第二阶段所进行的定价博弈。东南角“房间”下的第二层有四个“房间”，分别表示了两家公司定价行为的四种组合。

第二层的树枝与“房间”都类似于博弈树的终点结，我们可以写出每一种情况的支付。这些支付都是由公司所获收益减去投资成本构成，其单位以 10 亿美元计。

首先考虑图 6—1 中通向西南角的树枝，它表示只有串音公司投资的情况。于是，如果它选择高价，会获得收益 240 亿美元，减去 100 亿美元投资成本，

获得净利润 140 亿美元，记为获得支付 14；如果它选择低价，会获得收益 160 亿美元，减去 100 亿美元投资成本，获得净利润 60 亿美元，记为获得支付 6。此时，正如在“树房”第一层的西南角“房间”里所示，全球对话公司所获得的支付为 0。依此类推，可得到只有全球对话公司投资时的支付，如图 6—1 中东北角所示。此时，第一阶段博弈表中东北角“房间”里表示出了串音公司所获得的支付为 0。

如果两家公司在第一阶段都选择了投资，它们就要进行图 6—1 中东南角所示的第二阶段定价博弈。如果双方在第二阶段都选择了高价，它们将分别获得 120 亿美元收益，减去 100 亿美元投资成本，从而将分别获得 20 亿美元净利润，记为获得支付 2；如果双方在第二阶段都选择了低价，它们将分别获得 80 亿美元收益，减去 100 亿美元投资成本，从而将分别亏损 20 亿美元，记为获得支付－2；最后，如果一家公司选择了高价，另一家公司选择了低价，低价企业将获得收益 160 亿美元，即获得支付 6，而高价企业则获得收益 0，即获得支付－10。

同第 3 章中的多阶段博弈一样，我们必须采用逆推法解决这一博弈，即从第二阶段博弈开始。在两家独立公司定价决策问题中，我们立即会看到高价格策略会获得更多的支付。

第二阶段的定价博弈可以用第 4 章所提到的方法来解决。然而，非常明显，这个博弈是一个囚徒困境。低价是双方的占优策略，因此博弈结果将是第二阶段博弈表中的东南角房间，每家公司获得支付－2。[2]随后，我们将更详细地解释这些支付就是第二阶段博弈的均衡支付。

逆推法告诉我们，第一阶段的行动策略必须从第二阶段的均衡与支付结果推断得到。因此，我们能够把刚才计算得到的支付填进“树房”底层中的空“房间”与“半空房间”。这使得“树房”的底层有了已知的支付，如表 6—1 所示。

表 6—1　第一阶段投资博弈（代入第二阶段均衡的逆推支付）

		全球对话	
		不投资	投资
串音	不投资	0，0	0，14
	投资	14，0	－2，－2

现在我们能够用第 4 章中的方法解决这一同时行动博弈。你应该立即意识到表 6—1 中的博弈是一个懦夫博弈。它有两个纳什均衡，每一个都需要一家公司选择投资，另一家公司选择不投资；投资的公司获得巨大的利润。因此，每家公司都希望这个博弈均衡是自己投资而另一公司不投资。在第 4 章中，我们简略地讨论了使得两个均衡中某特定的一个被选择的因素，我们还指出存在每家公司都力图获得自己想要的结果但最终却是双方都选择投资而面临亏损的可能。事实上，后者更可能在现实生活的博弈中发生。在第 8 章中，我们会对这种博弈做进

一步研究，表明在混合策略中存在第三种纳什均衡。

对表 6—1 的分析表明在我们的例子中的第一阶段博弈不存在唯一的纳什均衡。这个问题并不严重，因为就像在前面章节中所做的那样，我们可以把这个不明确的解先搁置起来。如果第二阶段博弈没有一个唯一的纳什均衡，事情会变得更糟糕。因此，详细说明结果的推导过程非常重要，因为我们需要据以计算出第二阶段的支付并从而逆推回第一阶段。

在图 6—1 右下方所示的定价博弈是上述完整的两阶段博弈的一部分。然而，它本身也是一个具备参与人、策略与支付的完备自足的博弈。为了使这种双重特性更加明显，我们把它叫做一个完整博弈的**子博弈**（subgame）。

更一般化，子博弈是一个从原始博弈中某个特定结点开始的多次行动博弈的一部分。同样，子博弈树是完整博弈树的一部分，完整博弈树的结点就是子博弈树的根或初始结。一个多次行动博弈有多少决策结，就有多少子博弈。

6.1.2 多阶段博弈的结构

在图 6—1 所示的多层次博弈中，每一阶段都包含一个序贯行动博弈。然而，情况并不都是这样，许多时候，同时行动与序贯行动相互混合交织。我们再举两个例子来说明这一点，并进一步阐述在前面章节中所介绍的思想。

在第一个例子中，我们对串音—全球对话（CrossTalk—GlobalDialog）博弈稍加修改。假定其中一个公司——不妨设为全球对话公司，已经对光纤网络产品进行了 100 亿美元的投资。串音公司已了解这一情况并决定是否投资。如果串音公司决定不投资，全球对话公司就将获得单一定价权。如果串音公司决定投资，两家公司就将进行前面描述的第二阶段定价博弈。这种多阶段博弈的博弈树在其初始结上具有传统的树枝，在这些传统树枝所通向的结点中开始了一个同时行动子博弈。完整的博弈树如图 6—2 所示。

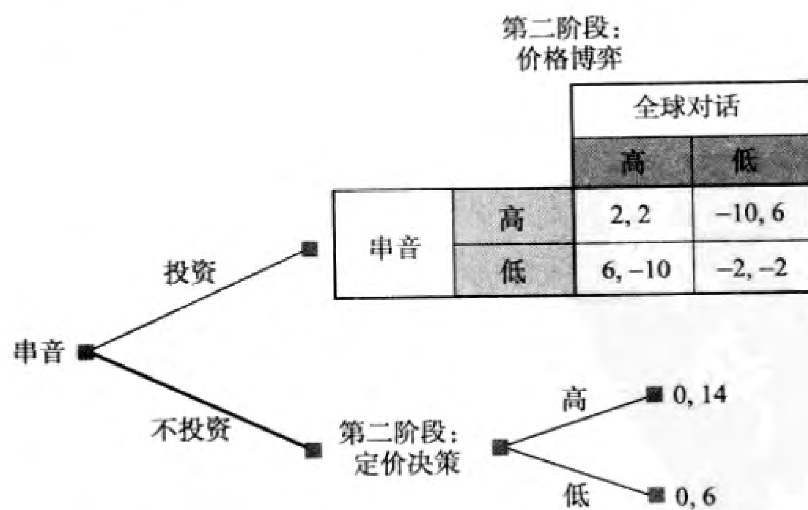


图 6—2 一个公司已投资情况下的两阶段博弈

当博弈树建立之后，就很容易对博弈进行分析。我们采用从第二阶段博弈中得到的均衡支付与串音公司第一阶段决策的一个更稠密的树枝来阐述在图6—2中用到的逆推分析法。串音公司意识到，如果它投资，就会导致囚徒困境，使得自己获得支付－2；如果不投资则会获得支付0。于是它会选择后者。串音公司关注的是最大化自己的收益而不是故意损害全球对话公司的利益，因此，串音公司是不会投资的，从而全球对话公司获得支付14。

上述分析提到了这种可能性，即全球对话公司为了能够保证从整个博弈中获得最大收益，于是试图在串音公司做出投资决定之前尽快进行投资。而串音公司也会选择同样方式来打击全球对话公司。在第10章中，我们将会研究能够使参与人确保这种优势的方法，即策略行动。

第二个例子源自橄榄球比赛。在每次开始之前，进攻方的教练都会为他的队伍选择进攻方式；同时，防守方的教练也会指示队伍如何布阵以抵抗进攻。而且，这些行动是同时进行的。假定进攻方有两种选择——安全进攻与冒险进攻，防守方可以有相应方法进行抵抗。如果进攻方计划采用冒险进攻方式，当四分卫看到敌人的防守阵形后能够在并列争球线上改变进攻方式。而防守方在得知这种改变之后也会相应改变防守阵形。于是，我们看到第一阶段是一个同时行动博弈，而这些行动组合之一又带来一个序贯行动子博弈。完整的博弈树如图6—3所示。

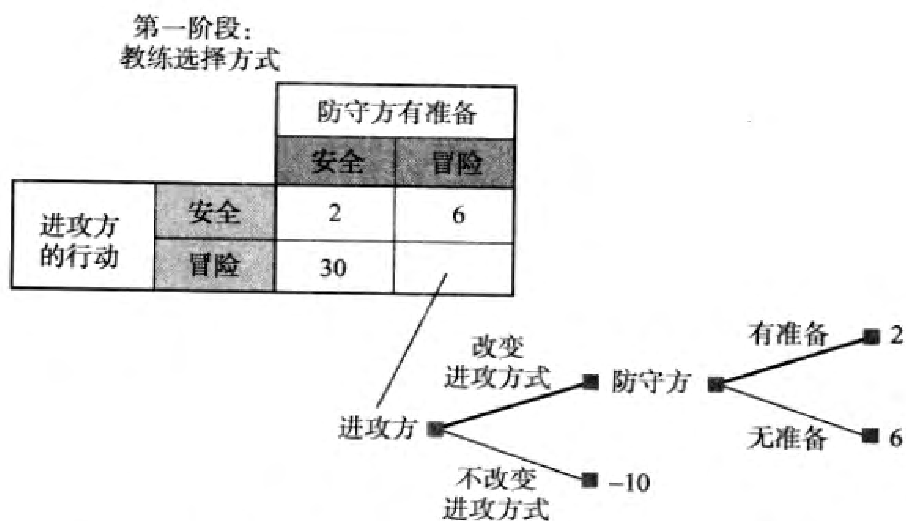


图6—3　序贯行动尾随其后的第一阶段同时行动

这是一个零和博弈，因此我们只表示出进攻方的支付，用进攻方期望获得的码数来表示。当防守方有准备时，安全进攻能够获得2码；当防守方无准备时，安全进攻也不会取得太好成绩，它能够获得6码。当防守方无准备时，冒险进攻能够获得30码；但当防守方有准备时，进攻方会倒退10码。如果此时进攻方不改变进攻方式，我们就在这个终点结上记支付为－10。如果此时进攻方改变了进攻方式（退回到安全进攻），当防守方有准备时支付为2，防守方无准备时支付为6。它与进攻方一开始就选择安全进攻时获得的支付是一样的。

图6—3中，我们用粗线表示序贯子博弈中被选择的分支。非常容易理解，

当进攻方改变进攻方式时，防守方一定会相应采取行动使进攻方得到 2 而不是 6。同时，进攻方为了能够得到 2 而不是 −10，也一定会改变进攻方式。逆推回去，我们就能够把支付 2 填进第一阶段同时行动博弈的右下方单元格中。于是，我们看到在纯策略中这个博弈没有纳什均衡。其原因与 4.8 节中网球博弈是一样的。一方（防守方）想要控制比赛（排好能够抵抗进攻方所选择进攻方式的防守阵形），而另一方（进攻方）则不想让比赛被控制（尽量使防守方排出错误阵形）。在第 7 章中，我们将介绍怎样计算这样一个博弈的混合策略均衡。其结果是进攻方会以 1/8 或 12.5%的概率选择冒险进攻。

6.2 在博弈中改变行动顺序

在前面章节中所提到的博弈要么是以序贯行动为本质的，要么是以同时行动为本质的。我们采用了合适的分析工具来预测每一类博弈的均衡。在本章第 1 节中，我们讨论了既有序贯行动又有同时行动的博弈。这种博弈需要同时运用两种工具来求解。但是，如果博弈既可以是序贯行动的又可以是同时行动的，情况会如何呢？改变博弈的形态，又改变分析的工具，会如何影响期望结果？

把序贯行动博弈变为同时行动博弈，只需要改变参与人做选择的时机或者可观察性。如果参与人在做决定前不能够观察到对手的行动，序贯行动博弈就变为了同时行动博弈。在这种情况下，我们将采用纳什均衡的方法来分析，而不采用逆推法来分析。相反，如果一个参与人能够在对方采取行动之前了解对方的行动，同时行动博弈就变为序贯行动博弈了。

博弈规则的任何改变都会改变其结果。这里，我们将阐释由于博弈类型的改变带来的多种可能结果。

6.2.1 同时行动博弈变为序贯行动博弈

1. 结果不改变

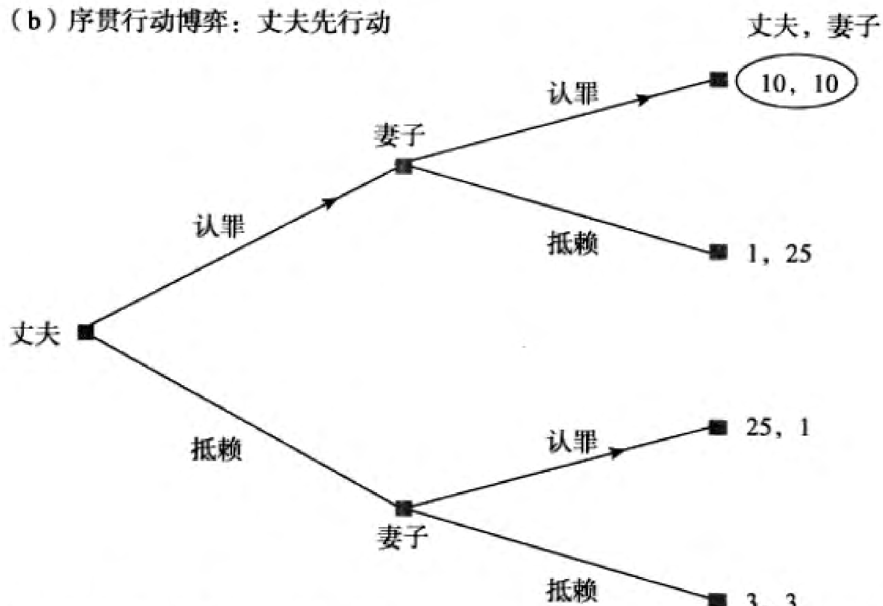
有些博弈无论其参与人是序贯行动还是同时行动，其均衡都是相同的，并且即使在序贯行动情况下，这一结果也与参与人的行动顺序无关。仅当双方（或所有方）都具有占优策略时，这一情形才会出现。我们可以在囚徒困境中看到这一点。

考虑在第 4 章中提到的丈夫和妻子均被审问的囚徒困境博弈，图 6—4（a）为复制的该博弈的同时行动版本，图 6—4（b）和图 6—4（c）为调整行动顺序后的序贯行动博弈。与表 4—4 一样，采用入狱时间（年）作为支付。因此，低数字比高数字更好。在图 6—4（b）中，丈夫在妻子之前做出决策，因此，妻子在做选择之前知道丈夫的决策；在图 6—4（c）中，情况正好相反。

(a) 同时行动

		妻子	
		认罪（背叛）	抵赖（合作）
丈夫	认罪（背叛）	10，10	1，25
	抵赖（合作）	25，1	3，3

(b) 序贯行动博弈：丈夫先行动



(c) 序贯行动博弈：妻子先行动

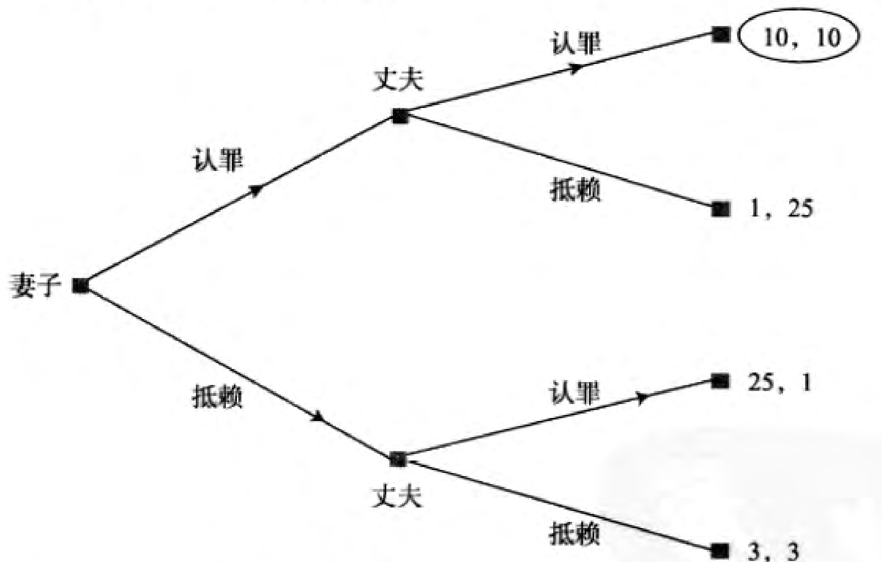


图 6—4　囚徒困境博弈的三个版本

图 6—4（a）中囚徒困境博弈的纳什均衡是双方都认罪（或者背叛与对方的合作）。如图 6—4（b）和图 6—4（c）所示，采用逆推法来求解序贯行动版本的该博弈。我们会看到，如果第一个人已经认罪，第二个人也会认罪（入狱 10 年总比入狱 25 年好）；如果第一个人抵赖，第二个人同样会选择认罪（入狱 1 年总比入狱 3 年好）。在已知第二个人的这种选择之后，第一个人一定会选择认罪（入狱 10 年总比入狱 25 年好）。均衡就是不管谁先做出决策，双方都入狱 10 年。

所以，这三个版本博弈的均衡是一样的。

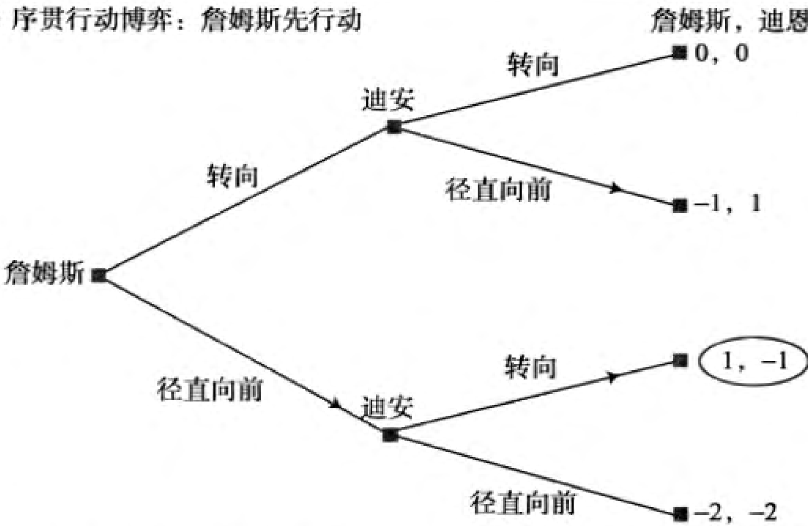
2. 先动优势

当博弈规则从同时行动变为序贯行动时，也可能会出现先动优势。如果同时行动时博弈有多重均衡，序贯行动则至少能够使先行者选择其想要的结果。我们运用懦夫博弈对此进行说明，在博弈中，两个青年驾车驶向对方，并都决定不转向。图 6—5（a）复制了表 4—14 中的策略式，图 6—5（b）和图 6—5（c）则表示了两个具有不同行动顺序的扩展形式。

（a）同时行动

		迪恩	
		转向（懦夫）	径直向前（勇士）
詹姆斯	转向（懦夫）	0，0	−1，1
	径直向前（勇士）	1，−1	−2，−2

（b）序贯行动博弈：詹姆斯先行动



（c）序贯行动博弈：迪恩先行动

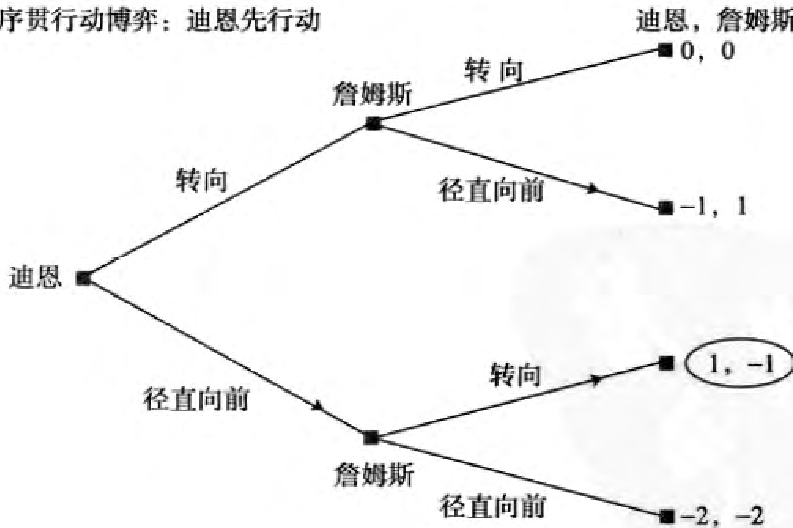


图 6—5　懦夫博弈的同时行动版本与序贯行动版本

在同时行动时，两个结果都是一人转向（懦夫）、一人径直向前（勇士），它们都是纯策略纳什均衡。如果不考虑某些特定的历史、文化与习俗，两者都不会

成为聚点。第 4 章中的分析认为协调行动能够帮助该博弈中的参与人，比如同意在两个均衡之间进行轮流。

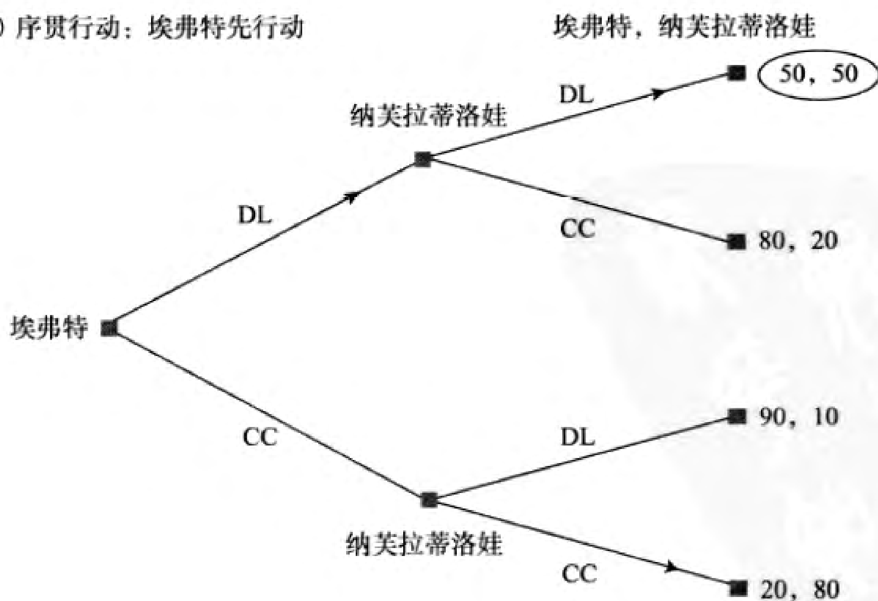
当我们改变这个博弈的规则，允许某人率先行动时，就不会存在两个均衡。更确切地说，我们会看到第二个行动者的均衡策略是选择与第一个行动者相反的行动。逆推法表明第一个行动者的策略是径直向前。图 6—5（b）与图 6—5（c）允许一个人先行动并且他的行动能够被观测到，这会导致一个唯一的逆推均衡，即先行动者获得支付 1，后行动者获得支付－1。在这种规则下，博弈中真实的行动已无关紧要，它可能会使得序贯行动下的该博弈非常无聊。尽管当规则改变以后，青年可能不会玩这种博弈，但它所导致的结果的改变是非常显著的。

3. 后动优势

在另外一些博弈中，当博弈规则从同时行动变为序贯行动时，也可能会出现后动优势。这可以用第 4 章中的网球博弈来说明。回想一下，在那个博弈中，纳芙拉蒂洛娃盘算回击的位置，而埃弗特则考虑在哪里防守。之前的版本假定双方都善于掩饰她们的意图，所以她们在最后一刻才同时行动。如果埃弗特不能掩饰她发球的意图，则纳芙拉蒂洛娃就可以做出相应反应从而变成博弈的后行动者，同样地，如果在埃弗特发球前纳芙拉蒂洛娃就靠向某一边了，则埃弗特就是后行动者。图 6—6 表示了这三种情况，图 6—6（a）复制了表 4—15 中同时行动的情形，图 6—6（b）与图 6—6（c）则表示了序贯行动时的两种不同行动顺序。

（a）同时行动

		纳芙拉蒂洛娃	
		DL	CC
埃弗特	DL	50	80
	CC	90	20



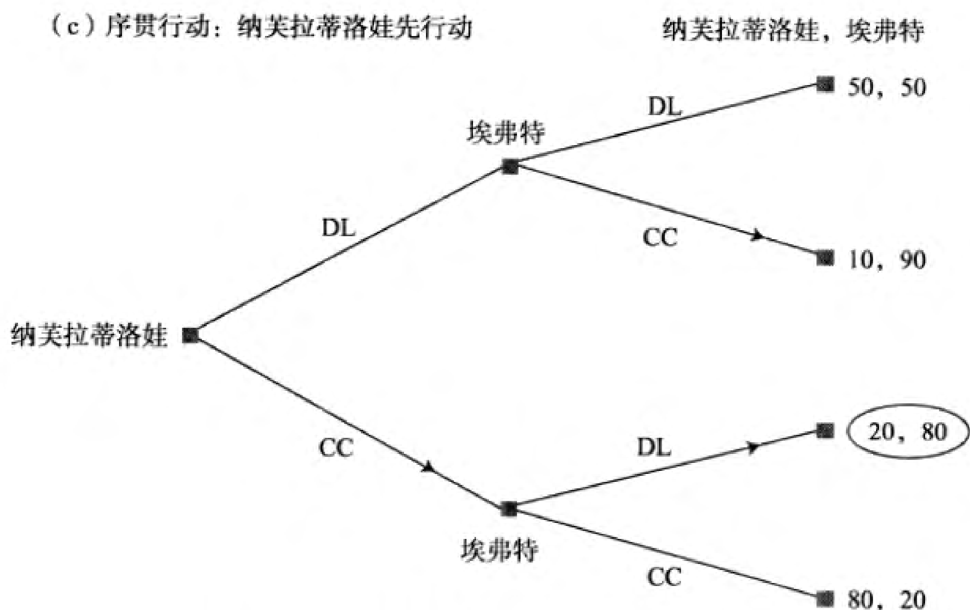


图 6—6　网球博弈的同时行动及序贯行动版本

这个博弈在同时行动的情况下不存在纯策略纳什均衡，而在不同行动顺序的序贯行动情况下，都存在唯一的逆推均衡，但均衡依谁先行动而有所不同。如果埃弗特先行动，纳芙拉蒂洛娃就会防守埃弗特选择的发球位置，从而埃弗特会选择底线球，在均衡中每人赢球的概率是一半对一半。如果顺序相反，埃弗特会选择把球发到纳芙拉蒂洛娃的防守空当，所以纳芙拉蒂洛娃应该防守底线，此时埃弗特赢球的概率是 80%。由于能够根据对手的行动做出最优反应，因此后行动者会具有优势。

第 7 章中我们会再次讨论该博弈的同时行动的情形，在那里我们知道存在一个混合策略纳什均衡。在该混合策略均衡中，埃弗特具有平均 62%的获胜概率，此概率高于其在序贯行动情形下率先行动时的 50%获胜概率，低于其后行动时 80%的获胜概率。

4. 双方都获益

当同时行动博弈变为序贯行动博弈时，我们非常直观地知道存在先动或后动优势，而更令人惊奇的是，在某些行动规则下，博弈双方都会比在其他规则下获得更高的技术。我们用联邦储备银行与国会之间关于货币政策与财政政策的博弈对此进行说明。在第 4 章中，我们研究了这个博弈的同时行动版本。图 6—7（a）复制了表 4—5 中的支付表，图6—7（b）与图 6—7（c）则表示两个序贯行动版本。为了更简要，我们用“平衡”与“赤字”代替国会的“预算平衡”与“预算赤字”，用“高”与“低”代替联邦储备银行的“高利率”与“低利率”。

（a）同时行动

		联邦储备银行	
		低利率	高利率
国会	预算平衡	3，4	1，3
	预算赤字	4，1	2，2

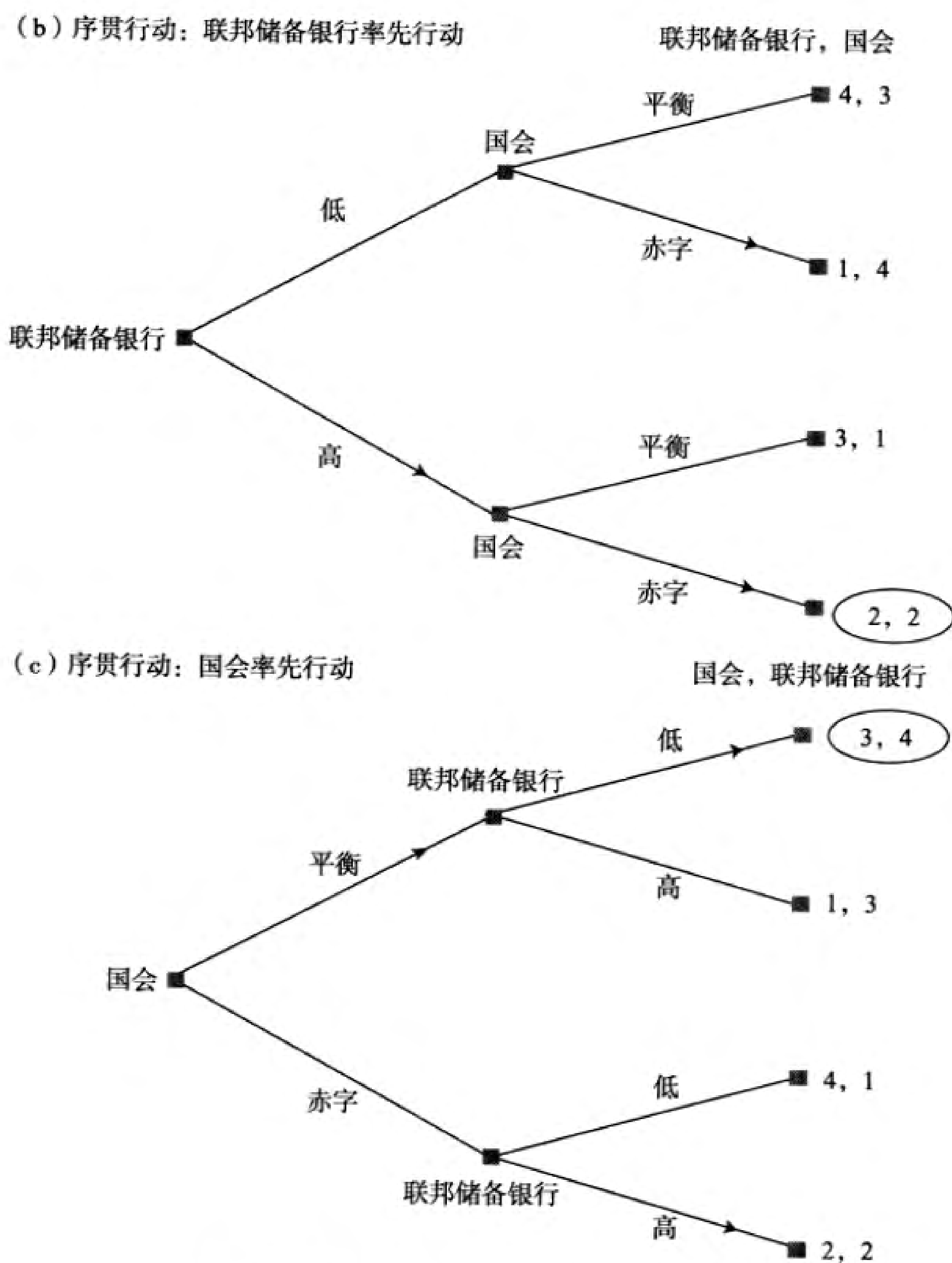


图 6—7 货币政策—财政政策博弈的三个版本

在同时行动情形下，国会具有占优策略（“赤字”），联邦储备银行了解这一点，于是会选择“高”，双方都获得支付 2。在序贯行动情形下，如果联邦储备银行先行动，会取得同样结果。联邦储备银行了解无论它采取哪种策略，国会都会选择“赤字”。于是“高”是联邦储备银行的更好选择，获得支付 2 而不是支付 1。

但是，在序贯行动情形下如果国会先行动，情况就会不同。现在，国会知道如果它选择“赤字”，联邦储备银行就会选择“高”；反之，如果它选择“平衡”，联邦储备银行就会选择“低”。这两种情况下，国会宁肯选择后者，此时它能够得到支付 3 而不是支付 2。于是这个行动顺序的逆推均衡就是国会选择“平衡”，联邦储备银行选择“低”。其结果是国会获得支付 3，联邦储备银行获得支付 4，都好于它们在另两种情形下所获得的支付。

由于图 6—7（c）中更好的结果来自于国会选择“平衡”，它不是图6—7（a）中的占优策略，因此这两个结果的差异更令人惊奇。为了解决这个明显的悖论，我们必须对占优的意义做更深的理解。“赤字”之所以能够成为国会的占优策略，是因为无论联邦储备银行选择哪种策略，国会选择“赤字”都比选择“平衡”好。由于此时国会必须在不知道联邦储备银行选择的情况下做出决策，在同时行动博弈中，这种在“赤字”与“平衡”之间的比较是非常重要的。国会必须考虑（或形成某种信念）联邦储备银行的行动并做出最优的反应。在我们的例子中，国会的最优反应总是“赤字”。如果国会后行动，由于它已经知道联邦储备银行的行动而仅需做出最优反应（总是“赤字”），此时的序贯行动博弈也会涉及占优的概念。然而，如果国会先行动，它就不能认为联邦储备银行的行动是既定的。相反，它必须意识到联邦储备银行的行动会受到自己第一步行动的影响。这里，国会知道联邦储备银行的反应将是用“高”对应“赤字”，用“低”对应“平衡”。国会会在这两者之间进行选择。由于受联邦储备银行的反应所限制，国会最偏好的结果“赤字”与“低”就不会出现了。

占优不再是一个同先行动者相关的概念，这种思想将在第 10 章中再次出现。在那里，我们将考虑这种可能性，即某一参与人可能故意改变博弈规则以成为先行动者。参与人能够用这种方法根据自己的偏好改变博弈的结果。

假定我们当前例子中的两个参与人能够选择在博弈中的行动顺序。于是，它们可能同意国会先行动。事实上，当面临财政赤字与通货膨胀时，在众多议员面前陈述的联邦储备银行主席经常会提出这种交易，他保证在国会紧缩财政之后降低利率。但口头上的承诺通常是不够的。因此，率先行动的技术性要求——既可以被后行动者观察到又是不可更改的——必须得到满足。在宏观经济政策中，非常幸运的是，美国财政政策的立法过程既透明又缓慢，而联邦储备银行的货币政策则变化迅速。所以这种国会先行动、联邦储备银行后行动的序贯行动博弈非常具有现实意义。

6.2.2 行动顺序上的其他变化

前面章节中列举了大量的博弈规则从同时行动变为序贯行动的例子。我们已经了解这种规则的改变如何以及为什么会影响一个博弈的结果。同样的例子也能够用来说明规则反方向变动的影响，即从序贯行动变为同时行动。于是，如果在序贯行动中存在先动优势或后动优势，在同时行动中就可能失去这种优势。同样，如果某种顺序能够使双方都受益，也可能会由于这种顺序的变更而使双方利益受损。

同样的例子也会告诉我们保持序贯行动性质不变时，行动顺序改变带来的影响。如果存在先动优势或后动优势，参与人从先行动变为后行动就可能因此受益或受损。同样，如果行动顺序关系到双方共同利益，源于外界因素的行动顺序的改变将使他们同时受益或受损。

6.3 分析方法的改变

博弈树是表示序贯行动博弈的常用方法，支付表则能够非常明了地表示同时行动博弈。然而，每一种技术也都能够适应另一种博弈方式。这里，我们展示如何将一种类型博弈中包含的信息转换到其他类型中去。在这个过程中，我们会提出一些将在接下来的分析中被证明有用的新思想。

6.3.1 用博弈树表示同时行动博弈

考虑第 4 章所提到的网球博弈。在那里，参与人行动非常迅速，以致可以认为二者同时行动，就像图 6—6（a）所示。但是，假定我们想用扩展式来表示这个博弈——比如用博弈树，图 6—8 将告诉我们如何做。

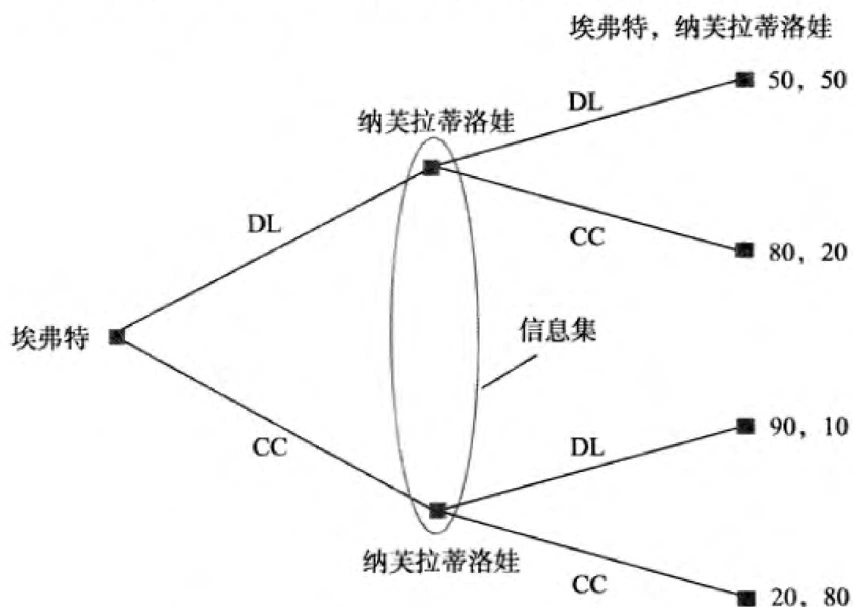


图 6—8　同时行动网球博弈的扩展形式

为了画出博弈树，我们必须选择一个参与人——比如埃弗特，将她的选择作为博弈树的初始结——她有两个选择的分支，DL 与 CC，以纳芙拉蒂洛娃的选择作为终点结。然而，由于行动基本是同时的，纳芙拉蒂洛娃必须在不知道埃弗特的行动的情况下进行选择。也就是说，她必须在不知道自己到底是在埃弗特的 DL 结点上，还是 CC 结点上的情况下进行选择。我们的博弈树必须采用某种方式表示出纳芙拉蒂洛娃的这种信息短缺状态。

我们用画一个椭圆圈住这两个相关结点的方法来说明纳芙拉蒂洛娃对自己所处结点的不确定性。（一个可选择的方式是采用虚线来连接；采用虚线是为了与博弈树分支的实线相区别。）椭圆内的结点被称为在此点行动的参与人的**信息集**

(information set)。这样一个集合表明参与人存在不完全信息。鉴于她可获得的信息有限，她无法对集合中的结点进行区分（因为在做出自己的决定之前，她不能观察到对手的行动）。因此，她在单个信息集上做出的策略选择就必须在该信息集所包含的所有结点上都指定同一个行动。也就是说，纳芙拉蒂洛娃在此信息集所包含的两个结点上要么选择 DL，要么选择 CC，而不能像在图 6—6（b）中那样在其中一个结点上选择 DL，在另一个上选择 CC。那里，博弈是序贯行动的，并且她后行动。

因此，我们必须修改关于策略的定义。在第 3 章中，我们把策略定义为行动的一个完全计划，它明确地说明在博弈规则确定该他行动的结点上，一个参与人所做出的行动。现在，我们重新准确地把策略定义为行动的一个完全计划，它明确地说明在博弈规则确定该他行动时的结点上的信息集里，一个参与人所做出的行动。

当参与人不确定的并非对手的行动而是博弈的外部环境时，也会涉及信息集的概念。例如，种植某作物的农场主可能对当前会影响其收成的气候如厄尔尼诺现象的威力不确定。此时，我们可以引入一个通常被称为"自然"的外部参与人，它率先行动并且在不同的可选项（这里是无厄尔尼诺现象、弱厄尔尼诺现象、强厄尔尼诺现象）中以源于我们的经验的概率做出选择。这样，我们就把这三个结点纳入了农场主的信息集中，他在这一信息集的约束下，针对这三个结点选择同一行动。

尽管这一概念表述起来很简单，但它并不能提供更简单的求解博弈的方法。因此，我们只在当它能够更简捷地传递某些意思时，才偶尔用它。在后面的第 9 章和第 14 章中，可以发现一些用以说明信息集概念的博弈案例。

6.3.2 用策略形式说明和分析序贯行动博弈

现在考虑图 6—7（c）所示的货币政策—财政政策序贯行动博弈，在博弈中国会首先行动。假定想要用规范式或策略式来表述它——也就是用支付表，表的行和列代表了两个参与人的策略。于是，我们首先需要说明具体的策略。

列出第一个行动者（国会）的行动策略非常容易。存在两种行动——平衡与赤字，国会也就有两种策略。对于第二个行动者，情况就变得更加复杂。记住，策略的定义为行动的一个完全计划，它明确地说明在博弈规则确定该他行动的每一个结点上，参与人所做出的行动。由于联邦储备银行在两个结点上行动（并且由于我们假定这个博弈事实上是序贯行动的，因此两个结点不能够被纳入一个信息集），并且在每一个结点上都能够选择"高"或"低"，因此它的选择方式就有四种组合。它们为：（1）如果"平衡"，就选"低"；如果"赤字"，则选"高"（为了简化，记为"如果 B 则 L，如果 D 则 H"）；（2）如果"平衡"，就选"高"；如果"赤字"，则选"低"（为了简化，记为"如果 B 则 H，如果 D 则 L"）；（3）总是低；（4）总是高。

我们把结果用表 6—2 所示的 2 行 4 列矩阵表示。最后两列与图 6—7（a）所

示的同时行动博弈的 2 行 2 列支付矩阵没有差异。这是因为如果联邦储备银行选择总是进行某项行动的策略，就仿佛联邦储备银行在行动时不会考虑国会已经采取了的行动，即仿佛它们的行动是同时的。但是计算前两列所列出的支付（此时联邦储备银行的行动取决于国会前一步的行动），就需要非常小心。

表 6—2　用策略式表示的货币政策与财政政策序贯行动博弈

		联邦储备银行			
		如果 B 则 L，如果 D 则 H	如果 B 则 H，如果 D 则 L	总是低	总是高
国会	平衡	3，4	1，3	3，4	1，3
	赤字	2，2	4，1	4，1	2，2

为了说明问题，首先考虑第二列第一行的单元格。此时，国会选择“平衡”，联邦储备银行选择“如果 B 则 H，如果 D 则 L ”。给定国会的选择，联邦储备银行在这个策略下的真实选择是“高”。因此，这个支付就是平衡和高的组合——国会获得支付 1，联邦储备银行获得支付 3。

通过逐个单元格检察，立即知道这个博弈有两个纯策略纳什均衡，我们用图中的灰色单元格加以表示。一个是左上方单元格，此时国会的策略是“平衡”，联邦储备银行的策略是“如果 B 则 L，如果是 D 则 H ”，因此它的真实选择是“低”。这个结果正好是序贯行动博弈的逆推均衡。在右下方单元格中存在另一个纳什均衡，此时国会选择“赤字”，联邦储备银行选择“总是高”。像所有的纳什均衡一样，博弈双方都没有明确的理由从导致这个结果的策略上偏离出去。国会如果转向“平衡”会获得更坏的结果，联邦储备银行也不能够通过转向其他三个策略而获得更好的结果，虽然它能这样做，正如选择“如果 B 则 L，如果 D 则 H”的情形一样。

当采用扩展式进行分析时，序贯行动博弈只有一个逆推均衡。但是当采用规范式或策略式进行分析时，却存在两个纳什均衡。为什么？

答案在于纳什分析与逆推分析逻辑上的不同性质。纳什均衡要求给定其中一方的策略后，另一方没有理由偏离。然而，逆推法并不认为后行动者的策略是给定的。相反，它要求存在行动机会时，采取最优的行动。

在我们的例子中，在存在行动机会时，联邦储备银行的“总是高”的策略并不满足最优的标准。如果国会选择“赤字”，“高”是联邦储备银行的最优反应。然而，如果国会选择“平衡”，而联邦储备银行必须做出反应，它将选择“低”，而不是“高”。因此，“总是高”并不是联邦储备银行在各种可能情况下的最优反应，也就不是一个逆推均衡。但是，纳什均衡的逻辑并不需要这样做，相反认为“总是高”是联邦储备银行的一个合理的策略。如果它这样做，“赤字”就是国会的最优反应。相反，当国会选择“赤字”时，“总是高”是联邦储备银行的最优反应（尽管它与“如果 B 则 L，如果 D 则 H”连在一起）。于是，尽管“赤字”与“总是高”这一对策略不是一个逆推均衡，但却是博弈双方的共同最优反应，构成了一个纳什均衡。

因此，我们可以认为逆推是一个更高的要求，是对纳什均衡的补充，能够帮助我们在多个纳什均衡中进行选择。换句话说，它是纳什均衡概念的一个精炼。

为了更精确地表达出这一思想，让我们回忆一下子博弈的概念。在整个博弈树的任何一个结点，我们都能够把从它开始的那部分博弈当成子博弈。事实上，当参与人连续地进行选择时，博弈的进程沿着连续的结点延伸，每一次延伸都可以被认为是一个子博弈的开始。采用逆推法得到的均衡符合在每一个子博弈上的特定连续选择，并得到了一条特定进行路径。当然，其他的进行路径也符合博弈规则。为了简化，我们把这些“其他的进行路径”称为**偏离均衡路径**（off-equilibrium paths），把由此产生的子博弈称为**偏离均衡子博弈**（off-equilibrium subgames）。

运用上述术语，我们可以说均衡路径是由参与人对选择另一不同行动结果的预期决定的——如果他们使博弈沿偏离均衡路径移动，从而开始了一个偏离均衡子博弈。逆推法要求所有参与人在博弈的每一个子博弈上都做出最优选择，而不管这个子博弈是否在最终均衡结果的路径上。

策略是完整的行动计划。因而，参与人的策略必须明确说明在每种可能情形下或者轮到他行动的每一个结点上——无论其是否在均衡路径上——采取的行动。到达某一结点时，与博弈下的进程相关的仅是始于该结点的行动计划，也就是适用于始于该结点的子博弈的那部分策略。这部分策略被称为子博弈策略的**连续**（continuation）。逆推法要求均衡策略的每一个子博弈的连续都是参与人在其相应结点处的最优反应，无论该结点和子博弈是否在均衡路径上。

回到国会首先行动的货币政策—财政政策博弈中，考虑其在策略式下的第二个纳什均衡。其进行路径是国会选择“赤字”，联邦储备银行选择“高”。在均衡路径上，“高”事实上是联邦储备银行应对“赤字”的最优反应。国会选择平衡则将成为一条偏离均衡路径的起点。它通向一个相当无足轻重的子博弈的起点——也就是联邦储备银行的决策。联邦储备银行所声称的“总是高”的均衡策略要求它在子博弈中选择“高”，但这不是最优的。这第二个均衡就是一个偏离均衡子博弈上的非最优选择。

相反，在表6—2左上角单元格中纳什均衡的行动均衡路径是国会选择“平衡”、联邦储备银行随即选择“低”，这是联邦储备银行在均衡路径上的最优反应。偏离均衡路径则是国会选择“赤字”，而联邦储备银行在给定策略“如果B则L，如果D则H”下，将选择“高”。联邦储备银行应对“赤字”的最优反应就是“高”，因此这个策略也是偏离均衡路径上的最优反应。

均衡策略要求在任何环境下策略的连续仍然是最优的。这样的要求是非常重要的，因为均衡路径本身就是参与人对自己做不同选择将带来的结果进行深思熟虑的结果。后行动的参与人为了获得好结果，喜欢威胁先行动者若他采取某些行动就将引起可怕的报复，或承诺先行动者只要他采取另一些行动就会有很好的报酬。但是先行动者总是怀疑这些威胁或承诺的**置信度**（credibility）。唯一打消这种疑虑的方法就是证明所声称的反应是最优的。如果这个反应不是最优的，威胁或承诺就不再可信，该反应也就不能在行动的均衡路径上被观察到。

采用逆推法得到的均衡称为**子博弈完美均衡**（SPE）。它是一组策略（行动

的完整计划)，这些策略对于每一个参与人而言，在博弈树的每一个结点，从这个结点出发的子博弈中的相同策略的连续对于在此采取行动的参与人都是最优的，而不管这些结点是否在行动的均衡路径上。更简单一点，一个 SPE 要求参与人选择的策略在博弈的每一个子博弈上都是纳什均衡。

事实上，由于在具有有限树与完美信息的博弈中，参与人能够观察到其他人之前所采取的每一步行动，因此在一个信息集中就不会有多个结点，逆推法也就能够找到唯一的子博弈完美均衡（一些异常例子除外)。考虑一下，如果你观察从最后一个行动者的最后一个决策开始的子博弈，参与人的最优选择就是能够获得最高支付的那一个策略，它正好是采用逆推法所选择的行动。当参与人沿着博弈树逆行时，逆推法排除所有没有理由的策略，其中包括不可置信的威胁与承诺。因此，按照这种方式得到的策略就是 SPE。于是，就本书宗旨而言，子博弈完美对于逆推法而言，是一个非常好的名称。在一些包含复杂信息结构与信息集的更高层次博弈理论中，子博弈完美就变为了一个含义更丰富的概念。

6.4 三人博弈

前面几节中，我们讨论的范围仅限于具有两个参与人、每个参与人有两个可选行动的博弈。但是这些方法也适用于更大与更一般的博弈。现在，我们用第 3 章中的街心花园博弈来对此进行说明。特别地，第一，我们将博弈规则从序贯行动改变为同时行动；第二，保持序贯行动不变但采用策略式进行表示与分析。首先，我们将图 3—6 中的序贯行动博弈树复制为图 6—9，并提醒你记住逆推均衡的概念。

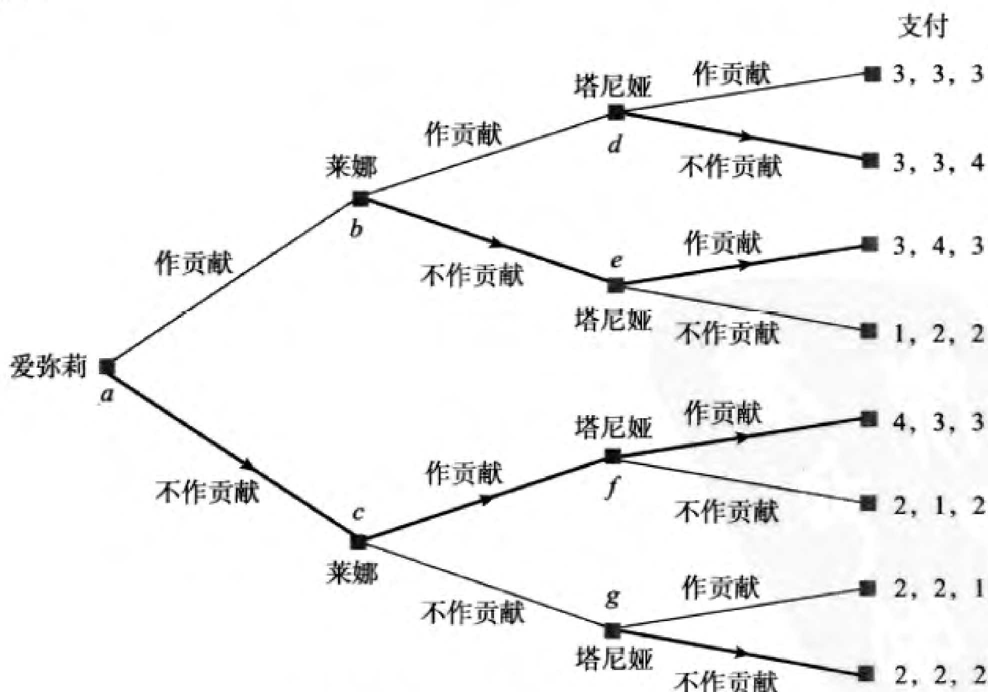


图 6—9　序贯行动的街心花园博弈

第一个行动者（爱弥莉）的均衡策略非常简单，就是“不作贡献”。第二个行动者（莱娜）则要从四个可能策略中进行选择（两个结点中的每一个都有两个选择），并选择策略“如果爱弥莉选择作贡献，就选择不作贡献；如果爱弥莉选择不作贡献，就选择作贡献”，或者简化一点，即“如果是C就选择D，如果是D就选择C”，甚至更简化一点，“DC”。塔尼娅则有16种可能的策略（四个结点，每个都有两种反应可选择），她的均衡策略是“如果爱弥莉选择C、莱娜选择C（即CC），就选择D；如果她们选择CD，就选择C；如果她们选择DC，就选择C；如果她们选择DD，就选择D”，或简称为“DCCD”。

同时，也要记住这些选择的理由。第一个行动者知道另外两人能够意识到除非她们都非常强烈地喜欢好花园并作出贡献，否则好花园将不会诞生，因此，她有机会选择不。

现在，我们把博弈规则改为同时行动（在第4章解决的那一个同时行动版本具有不一样的支付；这里，我们仍然保持与第3章中相同的支付）。支付矩阵如表6—3所示。通过逐个单元格检查，非常容易发现存在四个纳什均衡。

表6—3　　同时行动的街心花园博弈

塔尼娅选择：

作贡献

		莱娜	
		作贡献	不作贡献
爱弥莉	作贡献	3，3，3	3，4，3
	不作贡献	4，3，3	2，2，1

不作贡献

		莱娜	
		作贡献	不作贡献
爱弥莉	作贡献	3，3，4	1，2，3
	不作贡献	2，1，2	2，2，2

在同时行动博弈的其中三个纳什均衡中，两个人作贡献，而第三个人不作贡献。这些均衡与序贯行动博弈中的逆推均衡类似。事实上，每一个均衡都相当于一个有特定行动顺序的序贯行动博弈的逆推均衡。更进一步，在这个博弈的序贯行动版本中，任何给定的行动顺序都导致相同的同时行动支付表。

但是仍然存在第四个纳什均衡，即所有人都不作贡献。给定其中两个人特定的策略——不作贡献，另外一个人没有能力带来一个好花园，当然也就会选择不作贡献。于是，在序贯行动变为同时行动的过程中，先动优势就丧失了。在多个均衡中，只有一个使第一个行动者保留了高支付。

接下来，我们回到序贯行动版本——爱弥莉率先行动，莱娜第二，塔尼娅最后，但是采用规范式或策略式来表示它。在这个序贯行动博弈中，爱弥莉、莱娜、塔尼娅分别有2、4、16个纯策略。这意味着需要构造一个2×4×16的支付表。如果采用第4章中用到的构造具有三个参与人的支付表的方法，刻画这个博弈需要一个有16页2×4支付表的表格，它看起来非常杂乱。因此我们选择改组参与人。让塔尼娅为行的行动者，莱娜为列的行动者，爱弥莉则为页的行动者。于是，为了说明这个博弈，所需要的“全部”就是如表6—4所示的16×4×2博弈表。支付的顺序仍然与先前的惯例相同，按行、列、页的顺序加以记录。在本例中，意味着支付按照塔尼娅、莱娜、爱弥莉的顺序列在表中。

表 6—4　　策略式的街心花园博弈

	爱弥莉							
	作贡献				不作贡献			
	莱娜				莱娜			
塔尼娅	CC	CD	DC	DD	CC	CD	DC	DD
CCCC	3，3，3	3，3，3	3，4，3	3，4，3	3，3，4	1，2，2	3，3，4	1，2，2
CCCD	3，3，3	3，3，3	3，4，3	3，4，3	3，3，4	2，2，2	3，3，4	2，2，2
CCDC	3，3，3	3，3，3	3，4，3	3，4，3	2，1，2	1，2，2	2，1，2	1，2，2
CDCC	3，3，3	3，3，3	2，2，1	2，2，1	3，3，4	1，2，2	3，3，4	1，2，2
DCCC	4，3，3	4，3，3	3，4，3	3，4，3	3，3，4	1，2，2	3，3，4	1，2，2
CCDD	3，3，3	3，3，3	3，4，3	3，4，3	2，1，2	2，2，2	2，1，2	2，2，2
CDDC	3，3，3	3，3，3	2，2，1	2，2，1	2，1，2	1，2，2	2，1，2	1，2，2
DDCC	4，3，3	4，3，3	2，2，1	2，2，1	3，3，4	1，2，2	3，3，4	1，2，2
CDCD	3，3，3	3，3，3	2，2，1	2，2，1	3，3，4	2，2，2	3，3，4	2，2，2
DCDC	4，3，3	4，3，3	3，4，3	3，4，3	2，1，2	1，2，2	2，1，2	1，2，2
DCCD	4，3，3	4，3，3	3，4，3	3，4，3	3，3，4	2，2，2	3，3，4	2，2，2
CDDD	3，3，3	3，3，3	2，2，1	2，2，1	2，1，2	2，2，2	2，1，2	2，2，2
DCDD	4，3，3	4，3，3	3，4，3	3，4，3	2，1，2	2，2，2	2，1，2	2，2，2
DDCD	4，3，3	4，3，3	2，2，1	2，2，1	3，3，4	2，2，2	3，3，4	2，2，2
DDDC	4，3，3	4，3，3	2，2，1	2，2，1	2，1，2	1，2，2	2，1，2	1，2，2
DDDD	4，3，3	4，3，3	2，2，1	2，2，1	2，1，2	2，2，2	2，1，2	2，2，2

与联邦储备银行同国会之间的货币政策—财政政策博弈一样，在街心花园的同时行动博弈中也存在多个纳什均衡（在习题 6 中，我们要求你找出所有的纳什均衡）。但是根据在图 6—9 中找到的逆推均衡，只存在一个子博弈完美均衡。尽管逐个单元格检查能够找到所有的纳什均衡，但重复剔除劣策略会减少合理均衡的数目。这种剔除能够区别那些含有不可置信成分的策略（例如 6.3.2 节中联邦储备银行的“总是高”策略），因此它是有用的，能够自始至终将我们导向唯一的子博弈完美均衡。

在表 6—4 中，我们从塔尼娅开始，剔除她所有的劣（弱的）策略。这一步剔出了除表中第 11 行（DCCD）以外所有的策略，而 DCCD 已经被我们确定为塔尼娅的逆推均衡策略。对于莱娜仍然可以运用这一方法，此时我们必须将横跨表格中不同页的结果进行比较。例如，比较她的策略 CC 与 CD，我们就必须找到表格两页中关于 CC 的所有支付，将它与关于 CD 的支付进行比较。对于莱娜而言，重复剔除劣策略之后，只剩下策略 DC，而这正是她的逆推均衡策略。最后，爱弥莉就只需要比较与她的选择（“作贡献”与“不作贡献”）相联系的两个单元

格。她选择不作贡献会获得最高的支付，因此她会选择不作贡献。这样，我们就和前面一样找到了她的逆推均衡策略。

表6—4博弈表中唯一的子博弈完美均衡于是就与和每一个参与人逆推均衡策略相联系的单元格一致了。请注意，这里带给我们这个子博弈完美均衡的重复剔除劣策略的过程是通过将参与人按与真实顺序相反的顺序进行考虑来展开的。它符合逆推分析中所考虑的行动顺序，因此使我们能够精确地剔除掉那些对于每一个参与人而言不符合逆推法的策略。这样做，使我们剔除了所有不满足子博弈完美的纳什均衡。

6.5 总　　结

许多博弈包含多个组成部分，其中一些涉及同时行动，另一些涉及序贯行动。在两阶段（或多阶段）博弈中，“树房”能够用来说明这种博弈。它能够区分出行动的不同阶段以及将这些阶段联系起来的方式。由行动后来阶段产生的完整的博弈称为整个博弈的子博弈。

改变博弈规则从而改变行动的顺序可能会改变博弈的均衡结果，也可能毫无影响。同时行动博弈变为序贯行动博弈可能不会改变结果（如果博弈双方都具有占优策略），也可能会导致先动优势或后动优势，甚至可能会使博弈双方都变得更好。即使一个同时行动博弈没有均衡或存在多个均衡，它的序贯行动版本都通常只会有一个唯一的逆推均衡。类似地，一个拥有唯一逆推均衡的序贯行动博弈，它的同时行动版本可能会有几个纳什均衡。

当参与人不知道自己在哪一个特定节点上做决定时，同时行动博弈能够通过收集信息集中决策结的方式采用博弈树来说明。同样，序贯行动博弈也能够采用博弈表的形式来表示，此时，每个参与人的完全策略集合必须被仔细检验。求解策略式表示的序贯行动博弈时，可能会产生多个纳什均衡。潜在均衡的数量能够通过以置信度作为标准剔除某些可能的均衡策略的方法加以精简。这个过程能够带来序贯行动的子博弈完美均衡（SPE）。上述求解过程也同样适用于有多个参与人的博弈。

关键术语

连续（continuation）
偏离均衡子博弈（off-equilibrium subgames）
置信度（credibility）
子博弈（subgame）
信息集（information set）
子博弈完美均衡（subgame-perfect equilibrium，SPE）
偏离均衡路径（off-equilibrium paths）

习　题

1. 考虑第3章习题2图a所示的两人序贯行动博弈，将这个博弈用策略式（支付表）表示出来。找出这个博弈中的所有纯策略纳什均衡。如果存在多个均衡，说明哪一个是子博弈完美均衡，哪些不是，并说明理由（缺乏置信度的原因）。
2. 考虑第3章习题5的空中客车—波音博弈，以策略式表示博弈，找出所有纳什均衡。哪一个均衡是子博弈完美均衡？对那些不是子博弈完美的均衡，找出其缺乏置信度的原因。
3. 考虑有两个参与人A与B的博弈，A先行动选择“上”或“下”。如果A选择“上”，博弈就结束，每位参与人得到支付2；如果A选择“下”，则轮到B。B可选择“左”或“右”。如果B选择“左”，双方得到0；如果B选择“右”，则A得到3而B得到1。
   (a) 画出该博弈的博弈树，并找出子博弈完美均衡。
   (b) 用策略式表示这个序贯行动博弈，找出所有的纳什均衡，并指出哪一个是子博弈完美均衡，哪一个不是。如果不是，解释原因。
   (c) 用什么方法可以从博弈的策略式中找出子博弈完美均衡？
4. 考虑可乐行业，可口可乐与百事可乐是两家主要公司（为了分析方便，忽略其他公司），市场规模是80亿美元。每家公司可以选择是否做广告，广告成本为10亿美元；如果一家公司做广告而另一家不做，则前者抢得所有市场；如果两家公司都做广告，则各占一半市场，并付出广告成本；如果两家公司都不做广告，也各占一半市场，但不用支付广告成本。
   (a) 画出博弈的支付表，并找出当两家公司同时行动时的纳什均衡。
   (b) 假定博弈序贯进行，画出可口可乐公司率先行动时该博弈的博弈树。
   (c) 在(a)与(b)的均衡中，从可口可乐与百事可乐的共同观点来看，哪一个是最佳的？这两家公司要怎样才会有更好的结果？
5. 沿着海滩有500个小朋友，100个为一群，分为5群（以A、B、C、D、E按顺序表示，并且连续等距离排列），两个冰淇淋摊贩同时决定在何处贩卖，他们必须选择其中某一群小朋友所在的位置。

   如果有一个摊贩在某一群，则这群的100个小朋友都会买他的冰淇淋；对于没有冰淇淋摊贩的群，100个小朋友里面会有50个愿意走到隔壁群的摊贩那里去买，只有20个愿意走到距离两群的摊贩那里购买，没有人愿意走到距离三群或更远的摊贩那里。冰淇淋融化很快，所以走路去买的人不能帮没去买的人代买。

   如果两个摊贩选择相同的群，每个会得到50%的冰淇淋市场；如果他们选择不同群，则小朋友会走到较近的摊贩，距离相等就会各占50%的市场。每个摊贩要想办法使销售量最大。

   (a) 建立该博弈的5×5的支付表。你可以根据下面所述的内容开始刻画支付

表，也可以用它们来核对计算的结果：

如果双方都选择 A，则每个摊贩贩卖得 85 单位。

如果第一个摊贩选择 B，而第二个选择 C，则第一个卖得 150，而第二个卖得 170。

如果第一个选择 E，而第二个选择 B，则第一个卖得 150，而第二个卖得 200。

(b) 尽可能地剔除劣策略。

(c) 在剩下的表中，找出所有的纯策略纳什均衡。

(d) 如果把博弈变为序贯行动博弈，此时让第一人首先选择位置，第二人随后选择，找出子博弈完美均衡导致的位置与销售量。行动顺序的改变是如何帮助参与人解决第 (c) 问中的协调问题的？

6. 本章第 4 节中的街心花园博弈，如果用策略式来表示它的序贯行动版本，就会有 16×4×2 的支付表，如表 6—4 所示。在这个表中可以找到多个纳什均衡。

(a) 逐个单元格检查，找出表 6—4 中的所有纳什均衡。

(b) 从你所找出的纳什均衡集合中找出子博弈完美均衡。其他的均衡结果看起来与子博弈完美均衡相同——它们使得三个参与人中每人都获得了与子博弈完美均衡时相同的支付，但是它们是不同策略组合的结果。请解释它发生的原因。描述在非子博弈完美均衡中所存在的置信度问题。

7. 回忆在第 3 章习题 8 中的购物中心选址博弈。这一三人序贯博弈有一个类似于表 6—4 所示的街心花园博弈的博弈树。

(a) 画出这一购物中心选址博弈的博弈树，然后用策略式表示这个博弈。

(b) 找出博弈中的所有纯策略纳什均衡。

(c) 用反复占优找出子博弈完美均衡。

(d) 现在假定这个博弈仍然为序贯行动，但拥有不同的行动顺序——巨人、太阳神、弗里达。画出博弈树并找出这个博弈的逆推均衡。这个均衡与该博弈的原始版本的均衡有何不同？为什么？

8. 习题 7 中分析的购物中心选址博弈的规则明确指出，当三家百货公司都申请城市购物中心中的空间时，最大的两家会得到适用的空间。博弈的原始版本也指明在申请购物中心空间时，百货公司是序贯行动的。

(a) 假定三家百货公司同时申请。画出这个版本博弈的支付表并找出所有的纳什均衡。哪一个均衡是子博弈完美均衡？为什么？

(b) 现在假定三家百货公司同时申请城市购物中心时，抽签决定哪两家被选中，所以每家百货公司都有相同的机会进入城市购物中心。采用这种方式，如果三家百货公司同时申请，则每家有三分之二的概率 (66.7%) 被选中。画出这一新规则下的同时行动博弈的支付表，找出所有的纳什均衡以及子博弈完美均衡。

(c) 将 (b) 中所得到的结果与 (a) 中所得到的均衡加以比较，你是否得到了相同的纳什均衡？为什么？在抽签决定谁被选中的情况下，哪个均衡是聚点均衡？

9. 两个法国贵族，谢瓦利尔·沙格兰、马奎斯·德·雷纳，进行决斗。每人有一把只有一颗子弹的手枪。他们相隔 10 步距离，同时走向对方，每次走一步。每走一步后，他们都有可能开枪。开枪时，射中对方的概率取决于他们间的距离。k 步之后射中的概率是 $k/5$，因此，射中的概率在第 1 步后为 0.2，5 步之后上升为 1，此时，双方恰好彼此直面对方。如果一个人射击但没有命中而另一个人还没有射击，此时即使没有子弹的那位面临的必定是死亡，决斗仍将继续。贵族的道德标准保证了这些规则能够被执行。如果被击毙，将得到支付－1；如果击毙对方，得到支付 1；如果双方都没有被击毙，分别得到支付 0。

   这是一个拥有 5 步序贯行动并且每一步中参与人都同时行动（射击或不射击）的博弈。找出这个博弈的所有逆推（子博弈完美）均衡。

   提示：从第 5 步开始，此时他们恰好彼此直面对方。在这一步，建立这个同时行动博弈的 2×2 支付表，找出它的纳什均衡。现在回溯至第 4 步，此时击中的概率为 4/5，建立这一步的2×2的同时行动博弈表格，正确地在相应的单元格里标明未来将要发生什么。例如，如果一个人射击但没有命中而另一个人没有射击，于是另一个人就会等到第 5 步才射击以确保命中，于是博弈将会进入下一步，而此时你已经找到了均衡。利用所有这些信息，写出第 4 步 2×2 表格中的支付，并找到这一步的纳什均衡。用同样的方法逆推出剩余的步骤，就可以找到整个博弈的纳什均衡。

10. 试列举出一个与习题 9 决斗博弈有相似结构的商业竞争案例。

【注释】

[1] 有时，运用第 7 章、第 8 章的工具可找出整个博弈的同时行动部分的混合策略均衡，我们在本章相关处提及这种可能性并在后续章节的习题中练习这些方法。

[2] 就像囚徒困境中常发生的一样，如果企业之间能成功合谋采取高价，那么每个企业都可以获得较高的支付 2，但这样的结果不是一个均衡，因为每个企业都有动机欺骗对方从而得到更高的支付 6。

第 7 章 采用混合策略的同时行动博弈(Ⅰ)：零和博弈

在先前的第 4 章对同时行动博弈的研究中，我们遇到了采用那里描述的解决方法会无法求解的一类博弈，事实上，这一类博弈没有纯策略的纳什均衡。为了预测这一类博弈的结果，我们需要将策略与均衡的概念扩展到随机行动的范围内。这是本章与下一章的重点。

当一个参与人偏好某类行动，而其对手偏好避开这类行动时，随机行动就会出现。在零和博弈中就存在着这种参与人之间的直接冲突。在本章中，我们将详细讨论这类博弈。在军事冲突及大多数体育比赛等博弈中，随机行动概念的引入可能既有利于参与人，也便于解决无纯策略纳什均衡的难题。

我们以第 4 章结尾处的网球博弈为例来阐述这一思想。这是一个零和博弈，两个网球选手的利益是直接冲突的。埃弗特想要把球击到不能被纳芙拉蒂洛娃防守到的地方——打底线球（DL）或打斜线球（CC），而纳芙拉蒂洛娃则希望在埃弗特击球落地的地方成功地进行防守。在第 4 章中，我们指出，这种情况下，双方都希望自己的行动是不可预测的——它可以通过随机或无规律的行动来实现。在本章与第 8 章中，我们将更加全面地阐述这些思想。

7.1 什么是混合策略?

当参与人选择无规律地行动时，他们就会在他们的纯策略中随机进行选择。

在网球博弈中，纳芙拉蒂洛娃与埃弗特在最初给定的两个纯策略（DL与CC）中进行选择。我们把这两个纯策略的一种随机混合称为一个混合策略。

这类混合策略在一个完整的连续范围内变化。在一个极端，选择DL的概率为1（即一定会选），同时也意味着CC不会被选择（概率为0），这个“混合”就正好是纯策略DL。在另一个极端，DL被选择的概率为0，而CC被选择的概率则为1，这个“混合”也正好是纯策略CC。在两者之间则是整个概率集合：DL被选择的概率为75%，CC被选择的概率为25%；或者两者被选择的概率均为50%；DL被选择的概率为1/3，CC被选择的概率为2/3，等等。[1]我们可以将一个特定参与人混合策略的所有可能范围用代数的形式加以表示——DL被选择的概率为p，CC被选择的概率为（$1-p$），这里p为0到1之间的实数。

一个混合策略的支付被定义为构成它们的纯策略的相应支付的概率加权平均。例如，在4.8节的网球博弈中，针对纳芙拉蒂洛娃的选择DL，埃弗特选择DL的支付为50，选择CC的支付为90。于是，针对纳芙拉蒂洛娃的选择DL，埃弗特的混合策略（0.75DL，0.25CC）的支付为$0.75\times50+0.25\times90=37.5+22.5=60$。这就是埃弗特这个特定混合策略的**期望支付**（expected payoff）。[2]

你或许已经注意到混合策略类似于第5章中所研究的连续策略。事实上，它们正是连续可变策略的一种特殊形式。随着混合成为可能，每个参与人现在能从其纯策略的所有可能混合中进行选择（正如我们所说，纯策略是一个极端特殊的例子）。

纳什均衡的概念能够非常容易地扩展到混合策略情形。纳什均衡被定义为这样一组混合策略：每个参与人的选择都是其对其他参与人混合策略的最优选择，这种选择能够使他获得最大的期望支付。最优反应的分析方法可用来寻找每一参与人最优的混合概率（先前所描绘的变量p），它是其他参与人混合概率的一个函数。这些函数关系可以在以两个参与人的混合概率为坐标轴的坐标系中，用曲线的形式表示出来，而纳什均衡就是两条最优反应曲线的交点。在一个博弈中引入混合策略解决了我们在面对纯策略时通常都会碰到的纳什均衡不存在的难题。纳什经典理论表明，在通常情况下（这个范围能够覆盖本书中所有博弈，甚至更宽广的范围），每个博弈都具有混合策略纳什均衡。

在最宽的层面，将混合策略引入我们的分析中并不意味着与第5章中的连续策略理论并没有什么不同。然而，混合策略的特殊情形确实可带来一些特殊的概念与方法，因而它值得独立研究。首先，混合策略的最优反应曲线有一些特殊的性质，使得混合策略均衡也具有同样特殊的性质。其次，在零和博弈中对混合策略均衡的解释也与在非零和博弈中不一样。再次，起初我们将纳什均衡视为对手选择的自我确定（即信念）以及在此信念下该参与人的最优反应；然而，在混合策略中这样的观点需要重新解释。最后，将混合策略付诸实践非常棘手。在审视所有这些潜在的问题时，关于混合与混合策略均衡的经验证据应该被特别关注。在本章与下一章中，我们将讨论所有这些问题。在网球博弈的例子中，我们首先按部就班地开始，即用最优反应法找出混合策略纳什均衡，这样可启发我们体会到此博弈一开始就是使用混合策略的。

7.2 不确定行为：混合行动使对手难以预料

我们从 4.8 节中的网球博弈例子开始，这个例子中不存在纯策略纳什均衡。我们将展示向混合策略的扩展是如何弥补这一缺陷的，同时也将说明在由此产生的均衡中，每个参与人使其他参与人保持猜疑。

7.2.1 最优反应分析

混合策略是一种特殊的连续可变策略。于是要解决一个参与人采用混合策略的博弈，我们可用在连续策略博弈中寻找纳什均衡的方法，即最优反应分析。

首先我们需要描绘出对每个参与人有效的混合策略的整个范围。埃弗特有两个纯策略——DL 与 CC，按照规则她可以用任何概率对这两个纯策略进行组合。令代数变量 p 表示她选择 DL 的概率，于是 $(1-p)$ 就一定为她选择 CC 的概率，这里 p 的变化范围是从 0 到 1。p 的两个极端取值对应了两种特殊的纯策略情况：$p=1$ 意味着埃弗特只选择 DL；$p=0$ 则意味着埃弗特只选择 CC。我们将埃弗特的一般混合策略简称为埃弗特的 p—mix。类似地，纳芙拉蒂洛娃的一般混合策略为以概率 q 选择 DL，以概率$(1-q)$选择 CC，q 的变化范围从 0 到 1，简称为纳芙拉蒂洛娃的 q—mix。

我们将表 4—15 中的支付矩阵复制到表 7—1 中。尽管这是一个常和博弈，我们仍然分别表示两个参与人的支付，这样使我们能够更直观地考虑每一个参与人都试图为她们自己争取更高的支付。我们为每一个参与人都添加第三个策略，它代表两个纯策略的混合。因为不知道混合策略纳什均衡中参与人各自选择其纯策略的概率，我们分别用一般的数学符号 p 与 q 来表示。

表 7—1　　网球博弈中普通混合策略的期望支付

		纳芙拉蒂洛娃		
		DL	CC	q—mix
埃弗特	DL	50，50	80，20	$50q+80(1-q)$，$50q+20(1-q)$
	CC	90，10	20，80	$90q+20(1-q)$，$10q+80(1-q)$
	p—mix	$50p+90(1-p)$，$50p+10(1-p)$	$80p+20(1-p)$，$20p+80(1-p)$	

为了理解在 p—mix 行与 q—mix 列中所示的期望支付是如何计算得到的，考虑埃弗特选择 CC 而纳芙拉蒂洛娃选择 q—mix 的情形。当埃弗特选择 CC，如果纳芙拉蒂洛娃选择 DL，埃弗特的支付为 90；如果纳芙拉蒂洛娃选择 CC，埃弗

特的支付为 20。因此，当埃弗特选择 CC，而纳芙拉蒂洛娃采用以概率 q 选择 DL，以概率 $(1-q)$ 选择 CC 的混合策略时，埃弗特的期望支付为 $90\times q+20\times(1-q)$，这就是在单元格（CC，q－mix）中表示的埃弗特的支付。当埃弗特选择 CC，纳芙拉蒂洛娃选择 DL 与 CC 获得的支付分别为 10 和 80。于是当埃弗特选择 CC，而纳芙拉蒂洛娃选择 q－mix 获得的期望支付为 $10\times q+80\times(1-q)$，这就是在单元格（CC，q－mix）中表示的纳芙拉蒂洛娃的支付。同样的计算方法可应用于其他三个单元格，它们都是一个参与人采用混合策略，一个参与人采用纯策略的情形。当两者都采用混合策略时，单元格的数学表达会更复杂，但是我们并不需要它，因此让它为一个空格。

在均衡中，每个参与人都选择一个混合策略（p 或 q）作为她针对其他人的混合策略（q 或 p）的最优反应。像在第 5 章中一样，我们采取最优反应法可以很容易地找到均衡。纳芙拉蒂洛娃的最优反应规则必须使得对埃弗特的每一个 p 的取值，都有最优的 q 与之对应。同样，埃弗特的最优反应规则也必须使得对纳芙拉蒂洛娃的每一个 q 的取值，都有最优的 p 与之对应。我们把 p、q 都表示在横纵轴上，在图形中描绘出它们之间的最优反应关系，从而找出 p 与 q 的纳什均衡值（这里变量 p、q 的取值范围均从 0 到 1）。

首先从计算纳芙拉蒂洛娃的最优反应规则开始。我们构建一个图表，阐明纳芙拉蒂洛娃对埃弗特的每一个 p 的取值的可能反应，并从中找出每一种情况的最优反应。如图 7—1 所示，横坐标表示埃弗特的 p－mix 策略中 p 的取值范围（从 0 到 1），纵坐标表示纳芙拉蒂洛娃针对埃弗特的 p－mix 策略所采取的每一个纯策略的期望支付。

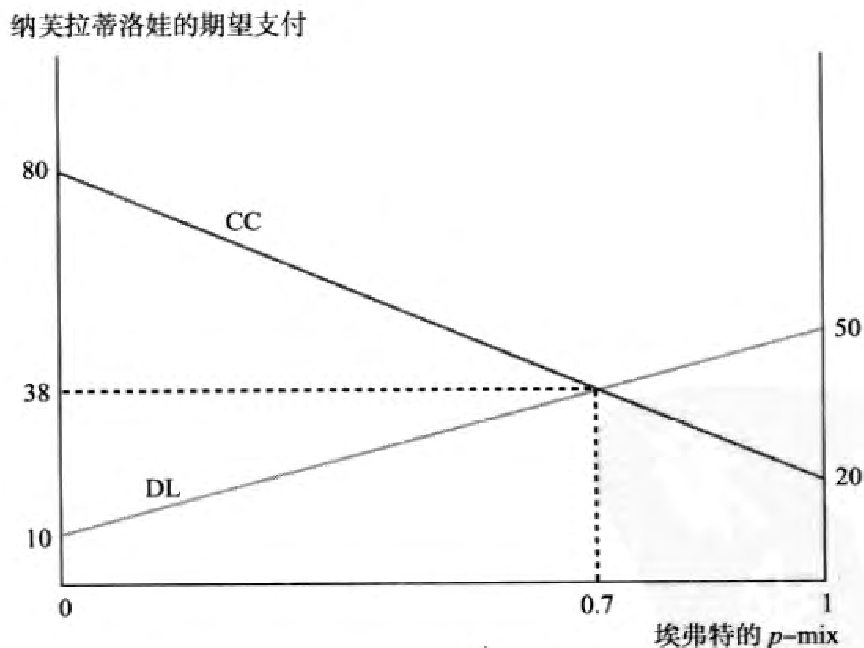


图 7—1　纳芙拉蒂洛娃最优反应的计算

纳芙拉蒂洛娃针对埃弗特的 p－mix 策略，选择 DL 的期望支付为 $50p+$

$10(1-p)$。这是一个线性函数，其图形为图7—1中上升的直线。它从 $p=0$ 时的取值10开始，终止于 $p=1$ 时的取值50。纳芙拉蒂洛娃针对埃弗特的 $p-$mix 策略，选择CC的期望支付为 $20p+80(1-q)$。这是图7—1中下降的直线，它从 $p=0$ 时取值80开始，终止于 $p=1$ 时的取值20。纳芙拉蒂洛娃从DL与CC的混合策略中获得的期望支付只不过是两种纯策略支付的一种平均，也就是说，它位于两者之间，或者说，它位于图7—1的两条期望支付线之间的某处。

计算出交点处的 p 值是非常有用的。此时 p 满足 $50p+10(1-p)=20p+80(1-p)$，即 $p=0.7$。在该点处的共同期望支付采用DL线计算为 $50\times0.7+10\times0.3=38$，或采用CC线计算也为 $20\times0.7+80\times0.3=38$。

我们想要找出纳芙拉蒂洛娃针对埃弗特的 p 的每一个可能取值的最优策略，这也正好是使她在每一种情况下都获得最高期望支付的策略。例如，对 $p=0$，图7—1表明纳芙拉蒂洛娃选择CC会比选择DL获得更高的支付。如果纳芙拉蒂洛娃选择CC与DL的某种混合，她的支付就是纯策略CC与DL支付的相应的平均。如果纯策略CC的支付高于DL的支付，当然也就会高于任何的这类平均。于是，纯策略CC对于纳芙拉蒂洛娃而言，不但优于DL，也是她的所有策略中的最优选择，无论纯策略还是混合策略。

对图中交点左边的所有 p 值，上述结论都是正确的。于是，对于图中交点左边的所有 p 值，纯策略CC对于纳芙拉蒂洛娃而言，针对埃弗特的 $p-$mix 策略，都优于DL，也优于她自己的任何混合策略。在这段 p 的取值范围中，纯策略CC是纳芙拉蒂洛娃的最优反应。同样，在交点的右边，纯策略DL是她的最优反应。[3]

针对埃弗特在交点处的 $p-$mix 策略，纳芙拉蒂洛娃选择CC与DL都是一样的（期望支付都是38）。且由于与CC与DL的一个混合相联系的期望支付位于直线CC与DL之间，因此她选取任何的 $q-$mix都是一样的。于是，纳芙拉蒂洛娃针对埃弗特在 $p=0.7$ 处的选择的最优反应为0到1之间的任何 q 值。

我们总结纳芙拉蒂洛娃的最优反应规则如下：

如果 $p<0.7$，选择纯策略CC（$q=1$）；
如果 $p=0.7$，0到1之间的任何 q 值都是最优反应；
如果 $p>0.7$，选择纯策略DL（$q=0$）。

为了了解这些规则背后可凭直觉触摸到的东西，注意到 p 为埃弗特将球击到底线的概率。纳芙拉蒂洛娃希望对埃弗特的击球有所准备。于是，如果 p 较低（埃弗特更可能打斜线球），纳芙拉蒂洛娃防守网前会更好；如果 p 较高（埃弗特更可能打底线球），纳芙拉蒂洛娃防守底线则会更好。这两者之间的关键点 p 的取值对于纳芙拉蒂洛娃而言，两种选择都同样好。这个点不是 $p=0.5$，因为不同的选择的结果（以各个不同混合策略的支付数值衡量）并不对称。

一个简单的代数计算可以深化基于图7—1的形象分析。针对埃弗特的 $p-$mix 策略，对于纳芙拉蒂洛娃而言，如果满足

$$20p+80(1-p)>50p+10(1-p),$$
或 $70(1-p)>30$，或 $70>100p$，或 $p<0.7$

则 CC 优于 DL（从而也优于任何的 q—mix）。如果相反的不等式成立（$p>0.7$），则 DL 优于 CC（也优于任何的 q—mix）。如果两个表达式相等（$p=0.7$），DL 与 CC（从而任何 q—mix）一样好。

迄今为止，我们所有分析针对的都是纳芙拉蒂洛娃的支付，这是由于考虑到如此开始会让人更易理解。但是，这是一个常和博弈，并且博弈矩阵通常只标明行参与人的支付。在这个博弈中，埃弗特是行参与人并且她的支付是我们平常在表中所见到的。我们能够通过埃弗特的支付找到纳芙拉蒂洛娃的最优反应吗？是的，我们恰好记得在常和博弈中，当埃弗特获得较低支付时，纳芙拉蒂洛娃会获得较高的支付。

图 7—2 如同图 7—1 一样，将埃弗特的 p—mix 策略放在横轴上，但将埃弗特的期望支付放在纵轴上。我们再次画出两条直线，一条对应纳芙拉蒂洛娃选择纯策略 CC，另一条对应她选择纯策略 DL。这里，直线 CC 从左向右上升，这对应于当埃弗特采用 DL 应对纳芙拉蒂洛娃的纯策略 CC 的概率越大（更高的 p）时，埃弗特的期望支付就会越高。类似地，直线 DL 从左向右下降，这是因为埃弗特采用 DL 应对纳芙拉蒂洛娃的纯策略 DL 的概率越大（更高的 p），埃弗特的期望支付就会越低。我们看到两条期望支付直线又在 $p=0.7$ 处相交。在该点的左端，直线 CC 低于直线 DL。于是，当 p 较低时，对于纳芙拉蒂洛娃而言，策略 CC 比策略 DL 更好，它使埃弗特获得更低的支付。在该点的右端，直线 DL 低于直线 CC。此时，对于纳芙拉蒂洛娃来说，策略 DL 优于策略 CC。在交点处，埃弗特的期望支付为 62，这正好等于 100 减去我们基于纳芙拉蒂洛娃的支付分析得到的她的支付 38，只不过分析基础有细微不同。你应该对零和博弈中的两种分析方法都感到满意。

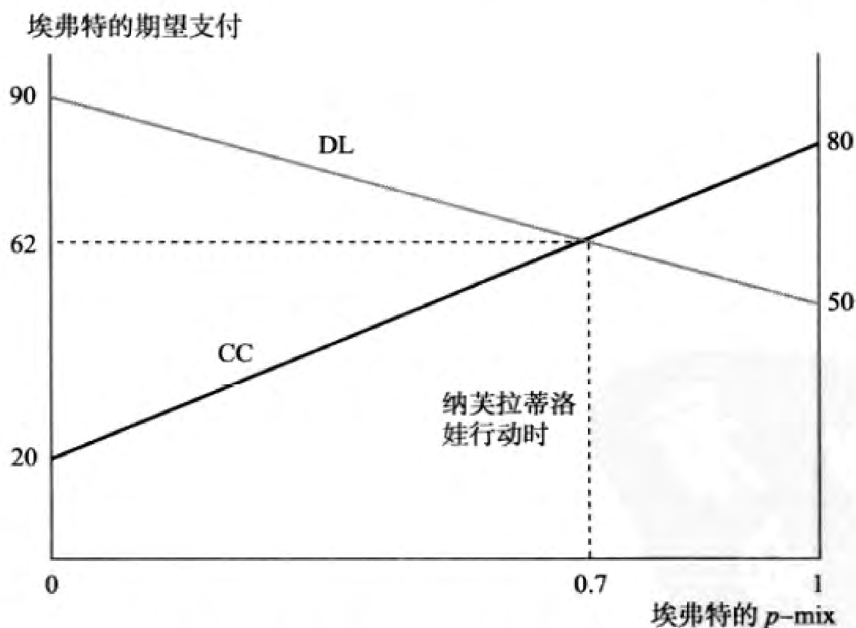


图 7—2　通过采用埃弗特的支付进行的纳芙拉蒂洛娃最优反应的计算

埃弗特的 q—mix 策略也能够用类似的方法得到，具体分析过程留给读者。其结果（埃弗特的最优反应规则）是

如果 $q<0.6$，选择纯策略 DL（$q=1$）；

如果 $q=0.6$，0 到 1 之间的任何 p 值都是最优反应；

如果 $q>0.6$，选择纯策略 CC（$q=0$）。

规则后面隐藏着一些可凭直觉触摸到的东西。注意到 q 为纳芙拉蒂洛娃防御底线的概率。埃弗特想要将球发到纳芙拉蒂洛娃未防御的地带。于是，当 q 较低时，埃弗特将球发到底线会更好；当 q 较高时，埃弗特将球发到网前会更好。

现在我们能够画出两条最优反应曲线，如图 7—3 所示。对于埃弗特选择 p，纳芙拉蒂洛娃做出的最优反应应是选择怎样的 q，这里 p 在横轴上（就像图 7—1 一样），q 在纵轴上。我们知道，对于 $p<0.7$，纳芙拉蒂洛娃选择纯策略 CC($q=0$) 会更好。她的最优反应曲线的这一段是沿着图形底边的水平实（灰）线。对于 $p>0.7$，纳芙拉蒂洛娃选择纯策略 DL（$q=1$）会更好。她的最优反应曲线的这一段是沿着图形顶边的水平实线。对于 $p=0.7$，纳芙拉蒂洛娃的所有选择都是一样的，无论纯策略还是混合策略，在 0 到 1 之间的任何 q 值都是最优反应，图中在 $p=0.7$ 处的垂直直线就是她对埃弗特的选择的最优反应。纳芙拉蒂洛娃的最优反应就为首尾相连的三条线段。正如我们在第 5 章中对普通连续策略所做的一样，我们将它称为纳芙拉蒂洛娃的最优反应曲线。其概念与第 5 章中的最优反应曲线相同，唯一的区别在于，对于混合策略，曲线有其特殊的形状。尽管它是由三条线段构成的，“曲线”显得有点用词不当，但是它是文章中的标准通用术语。

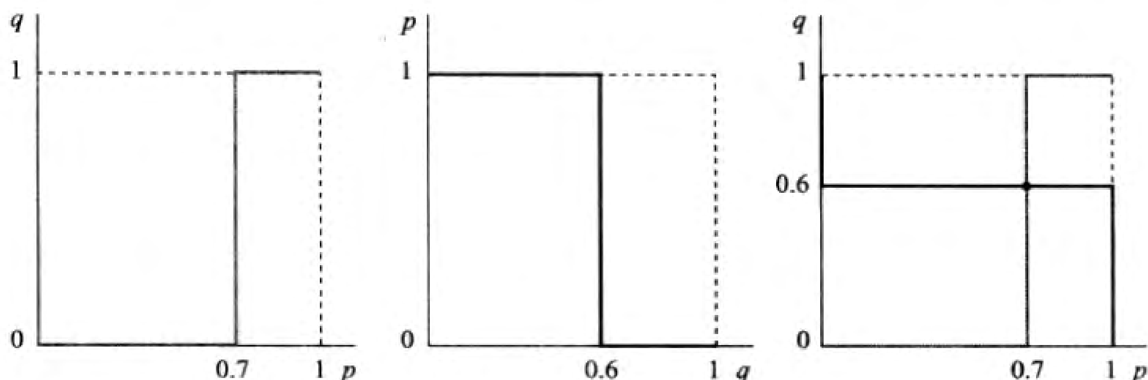


图 7—3　最优反应曲线和纳什均衡

图 7—3 的中间一幅图为埃弗特的最优反应曲线。这里，埃弗特的最优的 p 值由纳芙拉蒂洛娃所选择的 q 值决定，因此 q 为横轴，p 为纵轴。对于 $q<0.6$，埃弗特选择纯策略 DL($p=1$)会更好；对于$q>0.6$，埃弗特选择纯策略 CC($p=0$) 会更好；对于 $p=0.6$，任何选择都是一样的，无论纯策略还是混合策略。由三条首尾相连的线段构成的粗黑线就是埃弗特的最优反应曲线。

图 7—3 右边的一幅图为前两幅图的综合，它是将中间的图形沿对角线（45°线）映射，并将其添加到左边的那幅图上得到的，此时，p 为横轴，q 为纵轴。现在，两条曲线在 $p=0.7$，$q=0.6$ 处相交。这里，每一个参与人的混合选择相对其他选择都是最优的，因此，我们能够清楚地看到混合策略中纳什均衡的来历。

这幅图也表明了最优反应曲线不存在其他任何的交点。因此，本例中的混合

策略均衡就是该博弈的唯一的纳什均衡。更进一步，这个表述将纯策略视为对应于 p 与 q 的极值的特殊情况。因此，我们可以看到在正方形的边上，最优反应曲线没有交点，此时 p 与 q 等于 1 或 0。于是，我们有了另一种方法验证这个博弈不存在任何纯策略均衡。

可以观察到每一个参与人的最优反应都是用纯策略应对对方的几乎所有的混合策略。于是，当埃弗特选择 $p<0.7$ 时，纳芙拉蒂洛娃的最优反应为纯策略 CC，当埃弗特选择 $p>0.7$ 时，纳芙拉蒂洛娃的最优反应为纯策略 DL。只有在关键的 $p=0.7$，纳芙拉蒂洛娃的最优反应为混合策略，正如图 7—3 中的最优反应曲线的垂直部分所示。类似地，只有在纳芙拉蒂洛娃的混合策略的关键值 $q=0.6$ 处，埃弗特的最优反应是一个混合策略，即图 7—3 中她的最优反应曲线的水平部分。但是，这些看起来非常罕见的策略正是均衡中所显现的那一个。

p 与 q 的特殊值都有重要的性质。埃弗特的均衡 p 正位于纳芙拉蒂洛娃的最优反应曲线的垂直部分，此时，纳芙拉蒂洛娃的所有策略都是一样的，无论纯策略还是混合策略。纳芙拉蒂洛娃的均衡 q 则位于埃弗特的最优反应曲线的水平部分，此时，埃弗特的所有策略都是一样的，无论纯策略还是混合策略。于是每个参与人的混合均衡都使对方的所有策略无差异。我们将它称为**对手无差异性质**(opponent's indifference property)。

于是，这种最优反应曲线方法提供了对博弈的非常完整的分析。就像逐个单元格检查的方法（它检验了纯策略博弈的所有单元格）一样，最优反应曲线方法是一种当你想要找出一个博弈的所有均衡（不管是纯策略还是混合策略）时非常适用的方法。最优反应图能够将两种类型均衡表示在同一个地方。（它同样能够用于表示一个参与人采用混合策略，另一个参与人采用纯策略的均衡，尽管用图表示这种混合类型在一些非常例外的情况下会无能为力。）

如果你有理由相信一个博弈只有一个混合策略均衡，你就能够直接找出它，而不用烦琐地计算或图示最优反应曲线来找交点。利用“对手无差异性质”可建立一组方程。[4]埃弗特的均衡 $p=0.7$ 使纳芙拉蒂洛娃的采用 DL 与 CC 策略的期望支付相等。换句话说，p 是代数方程 $50p+10(1-p)=20p+80(1-p)$的解。同样地，纳芙拉蒂洛娃的均衡 $q=0.6$ 使埃弗特采用 DL 与 CC 策略的期望支付相等，q 是代数方程 $50q+80(1-q)=90p+20(1-q)$的解。在本章与第 8 章中，我们将介绍当一个参与人有三个或更多纯策略时这种方法的修正办法。

7.2.2 最小最大值法

由于这是一个零和博弈，因此当一个参与人是无差异的时，另一个也会这样。所以在这种特殊情形下，我们能够把对方的无差异看成参与人自己的无差异。在一个零和博弈中，一个参与人均衡的混合概率是给定此混合概率，对方选择任一纯策略都是无差异的（对对手而言）。这是因为在零和博弈中，如果对方有某个特定的纯策略能够带来正的收益，这个纯策略就一定是对参与人自己不利的。这里的直觉是在零和博弈中，参与人会慎重地混合自己的行动以使对方不能

通过提前预测而获利。我们通过将第 4 章中的最小最大值法扩展运用到处理混合策略中来发展这一思想。

在表 7—2 中，我们只列出了博弈的纯策略支付表，并且同在表 4—8 中一样，增加了行的最小值与列的最大值，埃弗特每一个策略的最小支付列在表格的最右端：DL 行为 50，CC 行为 20。这些最小值的最大值，或者叫埃弗特的最大最小值为 50。这是在考虑纳芙拉蒂洛娃会根据自己的最优利益进行反应时，埃弗特采用纯策略能够保证得到的最优值。为了获得这个支付（50）埃弗特会选择 DL。

表 7—2　　纯策略网球博弈的最小最大值分析

		纳芙拉蒂洛娃		
		DL	CC	
埃弗特	DL	50	80	min＝50
	CC	90	20	min＝20
		max＝90	max＝80	

在每一列的底部，我们列出了埃弗特获得成功时得到的最大收益。这些值代表了纳芙拉蒂洛娃在这一列中采用纯策略可能得到的最糟糕的支付。这些最大值的最小值，或者叫最小最大值是 80。这是在考虑埃弗特会根据自己的最优利益进行反应时，纳芙拉蒂洛娃采用纯策略能够保证得到的最优值。埃弗特若选择 CC 就会获得该支付。

两个参与人的最大最小值和最小最大值并不相等：埃弗特的最大最小成功收益为 50，低于纳芙拉蒂洛娃的最小最大值 80。正如我们在第 4 章中解释的，它表明该博弈没有纯策略均衡。在这个博弈中，每一个参与人都能通过选择一个适当的 DL 与 CC 的随机混合来获得更好的结果——埃弗特能够得到一个更高的最大最小值，纳芙拉蒂洛娃得到一个更低的最小最大值。

让我们把这一套参与人的策略集合扩展到包括随机行动或混合策略。我们这里从埃弗特的角度来进行分析，从纳芙拉蒂洛娃的角度进行分析也类似。首先，在表7—3中，较之表 7—2 中我们增加了一行埃弗特的 $p-$mix 策略，这里她以概率 p 选择 DL，以概率（$1-p$）选择 CC。我们列出针对纳芙拉蒂洛娃的每一个纯策略（DL 与 CC）埃弗特选择此混合策略的期望支付。$p-$mix 这一行的最小最大值就是这两个表达式中更低的一个。但是，这两个表达式哪一个更低依赖于 p 的取值。因此，我们必须画出这两条以 p 为参数的曲线，如图 7—4 所示。

表 7—3　　埃弗特混合策略网球博弈的最小最大值分析

		纳芙拉蒂洛娃		
		DL	CC	
埃弗特	DL	50	80	min＝50
	CC	90	20	min＝20
	$p-$mix	$50p+90(1-p)$	$80p+20(1-p)$	min＝?

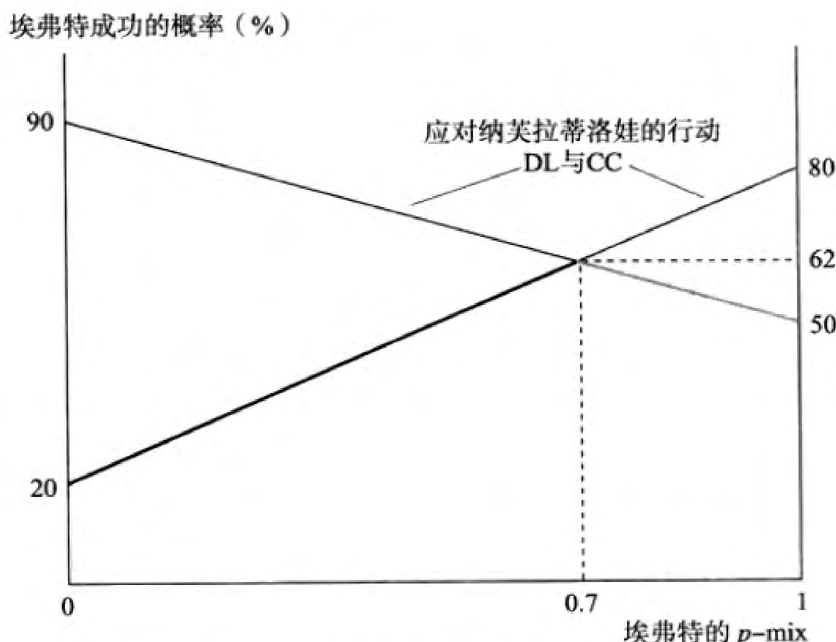


图 7—4　埃弗特关于 p 的最大最小选择

当纳芙拉蒂洛娃选择 DL 时，埃弗特采用 p－mix 策略获得期望支付 $50p+90(1-p)$，这是图 7—4 中向下倾斜的直线。当纳芙拉蒂洛娃选择 CC 时，埃弗特采用p－mix策略获得期望支付 $80p+20(1-p)$，这是图 7—4 中向上倾斜的直线。这两条直线相交，当

$$50p+90(1-p)=80p+20(1-p)$$
$$\text{或 } 70(1-p)=30p\text{，或 } 100p=70\text{，或 } p=0.7$$

将 $p=0.7$ 带入任何一个表达式，能够得到交点处的双方共同的期望支付为 62。

对每一个可能的 p－mix，埃弗特能够料到，在均衡中，纳芙拉蒂洛娃都会采取最有利于她自己的行动。由于这是一个零和博弈，最有利于纳芙拉蒂洛娃的行动就是最不利于埃弗特的行动。于是，埃弗特会料到当 $p<0.7$ 时，纳芙拉蒂洛娃会选择 CC；当 $p>0.7$ 时，纳芙拉蒂洛娃会选择 DL；如果埃弗特正好 70%选择下底、30%选择网前，纳芙拉蒂洛娃选择任何一种纯策略都是一样的。

在图 7—4 中，DL 与 CC 直线中较低的部分被加粗，以突出纳芙拉蒂洛娃迫使埃弗特在任何 p－mix 策略中获得的最低的成功概率。这条粗线（颠倒的 V 字形）表明了埃弗特在混合策略中选择任何 p 值能够得到的最低值。整个倒 V 字形就是表 7—3 中 p－mix 行最右端的最小值。事实上，这样做使我们能够忽略表 7—3 中对应于埃弗特的两个纯策略的最开始的两行。在最左边的一点，此时 $p=0$，即埃弗特选择纯策略 CC，此时她成功的概率为 20%。在最右边的一点，此时 $p=1$，对应于埃弗特的纯策略 DL，并获得 50%的成功。

更有趣的是，在 0 和 1 之间存在 p 值（代表了埃弗特真正的混合策略）使得埃弗特成功的可能超过 50%。这个结果说明了使纳芙拉蒂洛娃保持猜疑的好处。例如，如果 $p=0.6$，纳芙拉蒂洛娃会选择 CC。这一选择使埃弗特以 40%的概率选择 CC 从而获得 20%的成功机会；以 60%的概率选择 DL 从而获得 80%的成功

机会。埃弗特的**期望值**（expected value）为 $80\times0.6+20\times0.4=48+8=56$，它高于埃弗特选择两种纯策略中的任何一种时能够获得的最大最小值 50。于是，我们可以看到采用 60%DL、40%CC 的混合策略优于两种纯策略。但什么是最好的混合策略？

图 7—4 立即会告诉我们，埃弗特混合策略的最优选择位于倒 V 字形粗线的顶点。这个顶点代表了她选择任何 p 值获得的最小值中的最大值，并且它也是 DL 与 CC 直线的交点。我们已经计算出交点处的 p 为 0.7。于是，当埃弗特选择 p 为 0.7 时，不管纳芙拉蒂洛娃如何行动，她都有 62%的成功机会。其他任何的 p—mix 策略，纳芙拉蒂洛娃都能够选择相应的纯策略使得埃弗特的成功机会低于 62%，即沿着粗线的任何一边都会从顶点往下滑。于是，埃弗特选择 70% DL、30%CC 的混合策略是唯一不能够使纳芙拉蒂洛娃获利（对埃弗特不利）的策略。

总结一下，适当的混合策略使埃弗特的最大最小值从 50 提升到 62。从纳芙拉蒂洛娃角度展开的类似分析表明适当的混合策略（选择合适的 q）也能降低她的最小最大值。我们从表 7—1 可以看到纳芙拉蒂洛娃针对埃弗特选择 DL 的 q—mix 策略使埃弗特获得的期望支付为 $50q+80(1-q)$，而针对埃弗特选择 CC 的q—mix 策略使埃弗特获得的期望支付为 $90q+20(1-q)$。纳芙拉蒂洛娃的最小最大值位于两个式子相等处，即 $40q=60(1-q)$，或 $100q=60$，或 $q=0.6$。因此而使埃弗特的期望支付是 $50\times0.6+80\times0.4=30+32=62$。于是，纳芙拉蒂洛娃的最优混合策略能够使得她的最小最大值从 80 降到 62。当两个参与人的最优混合策略互相制衡时，埃弗特的最大最小值等于纳芙拉蒂洛娃的最小最大值，我们也就得到了一个混合策略的纳什均衡。

最优混合策略对应下的最大最小值和最小最大值相等，这是零和博弈的共有性质。此性质早在 1945 年就被冯·诺依曼（John von Neumann）和摩根斯坦（Oskar Morgenstern）证明了。但是更基本的概念在于通过观察即可发现混合策略可为每个参与人带来更好的结果。如果零和博弈没有纯策略纳什均衡，那么每个参与人最好别坚持每一纯策略，而让对方始终猜测你会使用怎样的策略。

7.3 纳什均衡作为一个信念与反应系统

当博弈中行动是同步的时候，参与人不能够对其他人的实际选择做出反应。事实上，每个人都会在考虑到其他人可能的选择的前提下，做出自己的最优选择。在第 4 章中，我们将这种考虑称为一个参与人关于其他人策略选择的信念。于是，我们将纳什均衡解释为当这种信念是正确的时候，每个参与人都选择其针对别人的实际行动的最优反应的一种形态。这个概念对于理解许多重要的博弈类型都是有用的，例如最特别的零和博弈、最小最大策略、占优、囚徒困境以及众多的合作博弈等。

然而，在第 4 章我们仅仅考虑了纯策略纳什均衡。于是，一个隐含的假设几

乎被忽略了，即每个参与人在他的信念中都确信其他人会选择一个纯策略。现在，我们考虑更一般的混合策略，信念的概念就需要做相应的重新解释。

参与人也许不能够确定其他人会采取什么行动。在第4章的协调博弈中，哈里希望能够碰到萨莉，他可能不能够确定萨莉会去星巴克还是本地咖啡店，但他的信念可能是萨莉会有50%的机会去任何一个地方。在网球博弈案例中，埃弗特会意识到纳芙拉蒂洛娃希望让她保持猜疑，于是她就不能够确定纳芙拉蒂洛娃可能的行动。

然而，对"不确定"与"错误信念"进行区分是很重要的。例如，在本章第2节的网球博弈案例中，纳芙拉蒂洛娃不能够确定埃弗特的选择，但她有关于埃弗特的混合策略的正确信念，即她知道埃弗特在两个纯策略之间选择的概率。有了关于混合策略的正确信念意味着能够知道或计算或猜测出其他参与人在基本的纯行动之间进行选择的概率。在我们例子的均衡中，结果是埃弗特的均衡混合策略为70%DL、30%CC。如果纳芙拉蒂洛娃相信埃弗特将会以70%的概率选择DL，以30%的概率选择CC，那么她的信念尽管不确定，但在均衡中却是正确的。

于是，我们通过另一种（数学意义上相同的）方式用信念来定义纳什均衡：每个参与人都形成了关于其他人选择的混合策略的概率的信念，并据此选择自己的最优反应。当这些信念是正确的时候，正如前边解释的，一个混合策略的纳什均衡就产生了。

在第8章中，我们将考虑非零和博弈中的混合策略及其纳什均衡。在那些博弈中，没有理由认为其他参与人追逐自己的利益会损害你的利益。于是，就不存在你需要对其他人隐藏你的企图，也就没有必要使其他人保持猜疑。然而，由于行动是同时发生的，每个参与人在主观上也许会对其他人采取的行动不确定，于是产生不确定信念，它又会使其对自己的行动不确定。这会导致混合策略均衡，而如何解释主观上不确定的正确信念就显得特别重要。

7.4 当一个参与人有三个或更多纯策略时的混合

以前对混合策略的讨论都局限于每个参与人都只有两个纯策略的情形。在许多情况下，每个参与人有更多纯策略，而我们应该计算出这些情况下的均衡的混合策略。然而，这些计算会迅速地变得非常复杂。对于真实的复杂博弈，我们将用计算机来寻找混合策略均衡。但是，对于一些小博弈，采用手工快速计算出均衡也是可能的。这个计算过程比通过计算机得到一个结果更能使我们理解均衡的作用原理。因此，在本节与下一节我们将解决一些相对较大的博弈。

这里，我们考虑这样一类零和博弈，在博弈中一个参与人只有两个纯策略，而其他参与人则有更多的纯策略。在这类博弈中，我们发现有三个（或更多）纯策略的参与人一般都只会在均衡中用两个。其他的策略在混合策略中不被采用，它们获得零概率。我们必须决定哪两个被采用或哪些不被采用。[5]

我们的例子是在网球博弈案例中，使埃弗特能够有第三个落点。除了底线球与斜线球，现在她可以考虑吊高球（Lob，更慢、更高、更长的回球）。均衡依赖于在纳芙拉蒂洛娃的两种防守策略下，埃弗特吊高球能获得的支付。我们从这个最容易举出的案例开始，之后再考虑一个一致或例外的案例。

7.4.1 一般情形

埃弗特现在有三个纯策略：DL、CC 与 Lob。我们让纳芙拉蒂洛娃有两个纯策略，防守 DL 或防守 CC。这个新博弈的支付表能够通过将表 7—1 增加一行（Lob）得到，如表 7—4 所示。现在由于你对常和博弈中的混合策略更加熟悉，我们就只表示出埃弗特的支付，并要求你记住纳芙拉蒂洛娃会选择能够使埃弗特获得更小期望支付的策略。

表 7—4　　带 Lob 的网球博弈的支付表

		纳芙拉蒂洛娃		
		DL	CC	q—mix
埃弗特	DL	50	80	$50q+80(1-q)$
	CC	90	20	$90q+20(1-q)$
	Lob	70	60	$70q+60(1-q)$
	p—mix	$50p_1+90p_2+70(1-p_1-p_2)$	$80p_1+20p_2+60(1-p_1-p_2)$	

表中前三行的支付非常简单。当埃弗特采用她的纯策略 DL 与 CC 时，她应对纳芙拉蒂洛娃的纯策略或 q—mix 策略的支付与表 7—1 一样，第三行也类似。当埃弗特选择 Lob 时，我们假定她应对纳芙拉蒂洛娃的 DL 与 CC 策略成功的概率分别为 70%与 60%。当纳芙拉蒂洛娃采用 q—mix 时，即以 q 的概率用 DL，以（$1-q$）的概率用 CC，埃弗特选择 Lob 的期望支付为 $70q+60(1-q)$，这就是埃弗特的 Lob 遇见纳芙拉蒂洛娃的 q—mix 策略的单元格所填的。

这里真正的新特性位于表中的最后一行。现在，埃弗特有三个纯策略，因此，她必须在她的混合策略中考虑三个不同的行动。混合策略不能够仅仅用数字 p 表示。更确切地，我们假定埃弗特以概率 p_1 选择 DL，以概率 p_2 选择 CC，留给 Lob 为剩下的概率 $1-p_1-p_2$。于是，我们需要两个数字 p_1 与 p_2 来定义埃弗特的 p—mix 策略。它们中的每一个作为概率，必须位于 0 与 1 之间。此外，其中两个相加不能超过 1。也就是，它必须满足条件 $p_1+p_2\leqslant 1$，因为用于 Lob 的概率 $1-p_1-p_2$ 必须非负。

采用埃弗特的 p—mix 策略的这个性质，我们能够得到当纳芙拉蒂洛娃选择纯策略 DL 时，埃弗特的期望支付为 $50p_1+90p_2+70(1-p_1-p_2)$，这就是在表 7—4 中最后一行第一个单元格中所填的。类似地，当纳芙拉蒂洛娃选择纯策略 CC 时，埃弗特的期望支付为 $80p_1+20p_2+60(1-p_1-p_2)$。我们不对用混合策略

应对混合策略得到的支付进行解释，这是因为它太长而且在以后的计算中并不需要。

技术上，在开始寻找一个混合策略均衡之前，我们应该证明不存在纯策略均衡。然而，这非常容易做到，因此我们把它留给读者而转向混合策略。

求解这种类型的常和博弈中的混合策略均衡的关键是从只有两个纯策略的参与人的角度采用最小最大值法。在这里，这个参与人就是纳芙拉蒂洛娃。这个方法有用是因为纳芙拉蒂洛娃的混合策略只用一个变量就能够定义，即用单一的概率 q 能够定义她的混合策略。在这个案例中，选择 DL 的概率能够完全定义她的混合策略。当 q 已知之后，选择 CC 的概率就是 $(1-q)$。

图 7—5 表示了纳芙拉蒂洛娃的 q 值在 0～1 间变化时，埃弗特选择 DL、CC 以及 Lob 的期望支付（成功的概率）。这些图形表示了表7—4中最右边一列。给定在零和博弈中的通常最坏情形，纳芙拉蒂洛娃对她的最小最大策略计算如下。对每一个 q，如果纳芙拉蒂洛娃在均衡中选择那一个 q—mix，埃弗特的最优反应就是选择使自己获得最高支付的策略。埃弗特的最优反应（也是对纳芙拉蒂洛娃而言的最坏方案）是在 q 这一点的三条线中最高的那一条。我们在图 7—5 中用粗线表示出这一系列最坏情形的结果。纳芙拉蒂洛娃所选择的最优的 q 使得埃弗特的支付在这一系列最坏情形结果中尽可能低（由此使自己的支付尽可能大）。

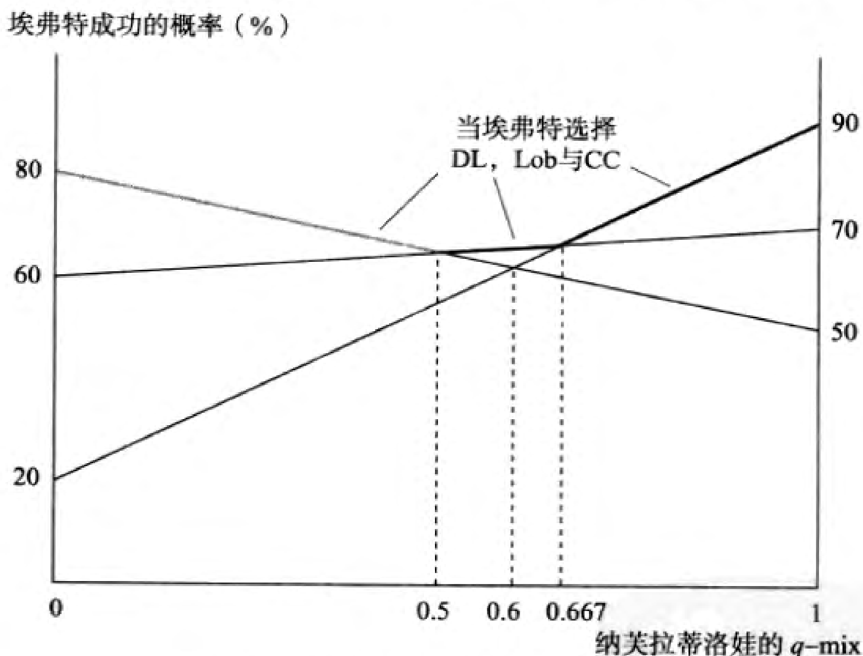


图 7—5　纳芙拉蒂洛娃的 q—mix 的图解

为了更精确地了解纳芙拉蒂洛娃的最优选择 q，我们必须计算出表示她的最坏情形结果的直线交点的坐标。最左端交点的 q 值使 DL 与 Lob 对于埃弗特是无差异的。这个 q 值必须使此时埃弗特采用 DL 与 Lob 策略时的支付相等。令两个式子相等，我们得到 $50q+80(1-q)=70q+60(1-q)$，即 $q=0.5$。第一个交点于是位于 $q=0.5$ 处，埃弗特在这一点的期望支付为 $50\times0.5+80\times0.5=65$。在第二个交点（最右边），采用 CC 与 Lob 对于埃弗特是无差异的。于是，这一点的 q 值必须使此时埃弗特采用 CC 与 Lob 策略时的支付相等。令 $90q+20(1-q)=$

$70q+60(1-q)$，得到$q=2/3$，最右边的点位于 $q\approx0.667$ 处，于是埃弗特在这一点的期望支付为 $90\times0.667+20\times0.333=66.67$。

现在，我们可以明确地描绘出针对纳芙拉蒂洛娃选择的不同的 q 值，埃弗特的最优反应。埃弗特的最优反应是：当 $q<0.5$ 时，选择 DL；当 $q>0.667$ 时，选择 CC；当 $0.5<q<0.667$ 时，选择 Lob。通常，埃弗特对大多数的 q 值的最优反应都是纯策略。当 $q=0.5$ 时，DL 与 Lob 对于埃弗特是无差异的，于是这两个纯策略以及它们的混合都是一样的。当 $q=0.667$ 时，CC 与 Lob 对于埃弗特是无差异的，于是这两个纯策略以及它们的混合都是一样的。

图 7—5 表明在所有的对纳芙拉蒂洛娃的最坏情形中，最好（坏处最小）的一个发生在左边交点，此时 $q=0.5$ 而埃弗特的期望支付为 65。粗线表示出了所有的最大值（对每一个 q），而这一点代表了它们中最小的一个，这就是纳芙拉蒂洛娃的最小最大值。在这一点，埃弗特得到根据纳芙拉蒂洛娃所选 q 采取最优反应能够获得的最低支付。于是，在 $q=0.5$ 时，纳芙拉蒂洛娃得到她最坏情形中支付的最大的一个，因此会在均衡中选择这个 q 值。

当纳芙拉蒂洛娃选择 $q=0.5$ 时，埃弗特在 DL 与 Lob 之间是无差异的，选择它们中的任何一个都能获得比选择 CC 更好的支付。于是，埃弗特在均衡中不会采用 CC，CC 在她均衡时的混合策略中将是无用的策略。

现在，我们可以像每个参与人只有两个纯策略的博弈一样继续均衡分析，纳芙拉蒂洛娃有两个纯策略 DL 与 CC，埃弗特有两个纯策略 DL 与 Lob。我们回到了熟悉的领域。于是，我们将计算过程留给读者，而只告诉结果。在博弈中，埃弗特最优的混合策略为用 0.25 的概率选择 DL，用 0.75 的概率选择 Lob。埃弗特选择这个混合策略应对纳芙拉蒂洛娃的 DL 与 CC 所获得的期望支付，分别为

$$50\times0.25+70\times0.75=80\times0.25+60\times0.75=65$$

这个支付就是埃弗特的最大最小值，它等于纳芙拉蒂洛娃的最小最大值，和通常的零和博弈混合策略纳什均衡的结果一致。

由于我们不能事先知道埃弗特的三个策略中哪一个不被采用，因此不能从 2×2博弈开始分析。但是，我们能够确信通常情况下会存在这样一个策略。一般情况下，这三条期望支付线会两两相交而不是交于一点。于是，这条上包络线就具有我们在图 7—5 中所看到的形状。它的最低点出现在其中两个策略的支付线的交点处，在这一点，第三个策略的支付线总处于其下。因此，参与人在这三个策略中进行选择时，总不会选择第三个策略。

7.4.2 例外情形

图 7—5 中三条线的位置与交点依赖于纯策略的支付。我们用一个特定博弈的支付来表示这些直线的一般结构。但是，如果这些参与人的关系特殊，我们就能够得到一些特殊的结构。

首先，如果埃弗特采用 Lob 应对纳芙拉蒂洛娃的 DL 与 CC 的支付是相等的，

Lob 的这一条直线就是水平的，整个范围内的 q 值都使纳芙拉蒂洛娃获得最小最大值。例如，在表 7—4 中的 Lob 行的两个支付都是 70，于是就很容易计算出修正后的图 7—5 的左边交点位于$q=1/3$处，右边交点位于 $q=5/7$处，埃弗特的最优反应就是 Lob。而我们则得到了一个特殊的均衡，即埃弗特采用一个纯策略而纳芙拉蒂洛娃采用混合策略。

其次，如果埃弗特采用 Lob 应对纳芙拉蒂洛娃的 DL 与 CC 的支付在右边部分低于图 7—5 中的位置（或者说其他两个策略在右边部分更高），所有三条直线就能够交于一点。例如，如果埃弗特采用 Lob 应对纳芙拉蒂洛娃的 DL 与 CC 的支付分别为 66 与 56，而不是 70 与 60，于是，在 $q=0.6$ 处，埃弗特采用 Lob 的期望支付变为 $66\times0.6+56\times0.4=39.6+22.6=62$，与采用 DL 和 CC 获得的支付相等。于是，埃弗特在 $q=0.6$ 处采用三个策略中的任何一个都是一样的，对它们进行任意混合也是一样的。

在这一特殊情形中，埃弗特的均衡的混合概率没有被完全决定。在这里，包括所有三个策略的整个混合范围都能够使纳芙拉蒂洛娃在 DL 与 CC 以及它们的混合之间是无差异的，于是愿意混合。然而，为了取得她的最小最大值，纳芙拉蒂洛娃必须采用 $q=0.6$ 的混合策略。如果她不这样做，埃弗特的最优反应将转向选择纯策略之一，如此将损害纳芙拉蒂洛娃的利益。我们不打算求出埃弗特均衡混合策略变动的精确范围，因为这个例子只是一个支付数字的特殊组合，因此相对不重要。

7.4.3 一个混合策略占优的情形

如果埃弗特采用 Lob 应对纳芙拉蒂洛娃的 DL 与 CC 的支付低于使三条线交于一点的取值，又会怎样？表 7—5 给出了这样一个支付矩阵。

表 7—5　　当 Lob 不再是一个最优反应的网球博弈

		纳芙拉蒂洛娃	
		DL	CC
埃弗特	DL	50	80
	CC	90	20
	Lob	75	30

当我们描绘埃弗特采用三个纯策略应对纳芙拉蒂洛娃所选 q 时，立即会发现 Lob 从图 7—5 中的位置向下移动了。图 7—6 表示了这些直线的新的结构，并决定出了纳芙拉蒂洛娃的最小最大值。使我们获得直线 DL、CC 与 Lob 的位置以及它们的交点的计算过程与以前相同，因此我们省略掉了这些细节。我们随后将对直线 Mix 进行解释。

很明显，根据这些数值，Lob 对于埃弗特而言不是一个好的策略。事实上，直线表明 Lob 的支付在任何时候都低于 DL 与 CC 之一，因此也低于这两条线的

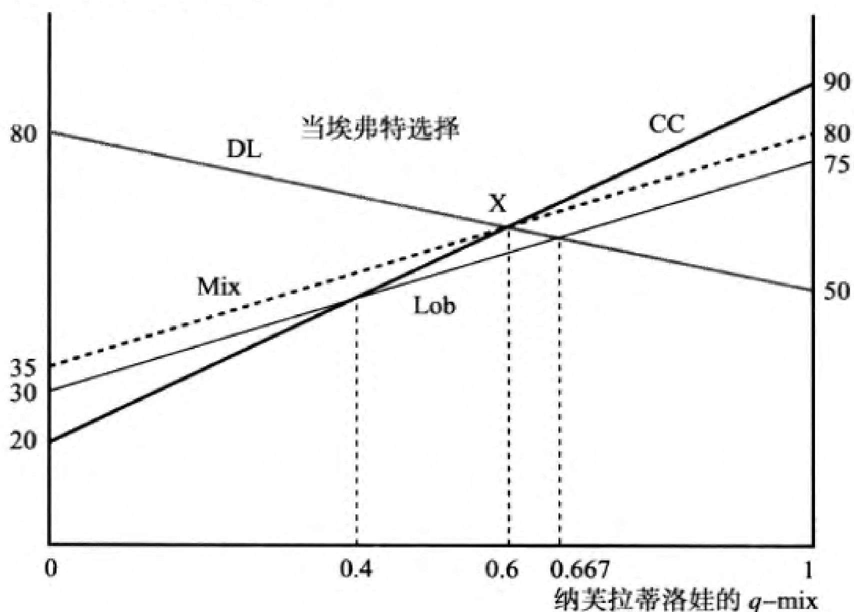


图 7—6 通过一个混合策略占优

上包络，同样也低于 DL 与 CC 直线的交点。于是，对纳芙拉蒂洛娃混合策略的每一个 q，都至少会有纯策略 DL 与 CC 中的一个使得埃弗特获得比采用 Lob 更高的支付。图 7—6 表明，如果 $q<0.667$，DL 对于埃弗特而言是一个比 Lob 更好的反应；如果$q>0.4$，CC 就好于 Lob；并且，当 $0.4<q<0.667$ 时，DL 与 CC 都好于 Lob。换句话说，Lob 对于埃弗特而言不再是最优反应。

然而，Lob 既不被 DL 占优也不被 CC 占优，因为当 $q<0.4$ 时，Lob 优于 CC，而当 $q>0.667$ 时，Lob 又优于 DL。于是，很明显，Lob 不能因为占优的理由而被从埃弗特的考虑中剔除。

但是，现在允许采用混合策略，我们能够考虑 DL 与 CC 的混合。这样一种混合能够占优 Lob 吗？答案是肯定的。考虑分别采用概率 p 与 $(1-p)$ 的 DL 与 CC 的一个混合策略。它对应于纳芙拉蒂洛娃的纯策略 DL 的支付为 $50p+90(1-p)$，它对应于纳芙拉蒂洛娃的纯策略 CC 的支付为 $80p+20(1-p)$。于是，它对应于纳芙拉蒂洛娃的 q—mix 策略的支付为$[50p+90(1-p)]q+[80p+20(1-p)](1-q)=[50q+80(1-q)]p+[90q+20(1-q)](1-p)$，这个表达式代表了直线 DL 与 CC 分别以比例 p 与 $(1-p)$ 进行的加权平均。

换句话说，埃弗特采用 p—mix 应对纳芙拉蒂洛娃的 q—mix 的期望支付是一条居于 DL 与 CC 中间的平均直线。自然，平均直线必将穿过 DL 与 CC 的交点。如果在某些 q 值上，两个纯策略 DL 与 CC 获得相同的支付，于是它们的任意混合也将得到这个支付。

在埃弗特的混合策略中，如果 $p=1$，她选择纯策略 DL，她的混合策略线也正好是 DL 直线；如果 $p=0$，混合策略线就是 CC 直线。当 p 在范围 0 到 1 之间变化时，p—mix 线便从 CC 直线旋转至 DL 直线，并总是经过 CC 直线和 DL 直线在 $q=0.6$ 处的交点。在所有的这些对应于不同的 p 值的直线中，有一条会平

行于直线 Lob。这条直线就是图 7—6 中标注为 Mix 的第四条直线（点状线）。因为直线 Lob 位于 DL 与 CC 交点的下方，于是它就完全低于这条代表 DL 与 CC 的特殊混合的平行直线。于是，无论 q 取任何值，Lob 获得的支付都低于这个特殊的混合策略。换句话说，这个混合策略占优于 Lob。

接下来，需要找到定义这个占优于 Lob 的混合策略的概率。从结构上看，这条混合策略直线平行于 Lob 并在 $q=0.6$ 处穿过交点，交点处期望支付为 62。于是，能够非常容易地计算出 Mix 直线的纵坐标。我们运用 Lob 的斜率 $=(75-30)/(1-0)=45$ 这一已知条件。于是，当 $q=0$ 时，Mix 线的纵坐标用方程 $45=(62-v)/(0.6-0)$ 中的 v 表示，得到 $v=35$；同样地，当 $q=1$ 时，解方程 $45=(v-62)/(1-0.6)$ 得到纵坐标 $v=80$。对于在 0 到 1 之间的 q 的任意取值，Mix 线的纵坐标为 $80\times q+35\times(1-q)$。现在，把这个结果与刚才推导出的通常的 $p-$mix 表达式相比较。由于两者应该相同，$p-$mix 表达式中多余的部分，$50p+90(1-p)$ 必须等于 80，而 $80p+20(1-p)$ 必须等于 35，皆可得出 $p=0.25$。

于是，埃弗特以 0.25 的概率选取 DL，以 0.75 的概率选取 CC 的混合策略得到了图 7—6 中的 Mix 线。无论纳芙拉蒂洛娃选择任何 q 值，这个混合策略都能使她获得比 Lob 更好的期望支付。换句话说，从埃弗特的角度出发，这个特定的混合策略占优于 Lob。[6]

现在，我们已经表明，如果一个策略是绝非最优反应，那么就能够发现一个混合策略占优于它。[7]在这个过程中，我们已经将占优概念的范围扩展到包括混合策略的占优。其逆定理也是正确的。如果一个策略被另一个策略占优，即使是一个混合策略，它也不再是对其他参与人策略的一个最优反应。于是，我们能够将第 4 章中提出的与占优相联系的其他一些概念也扩展到允许混合策略的情形。我们能够成功地剔除那些被别的策略（纯的或混合的）占优的策略（纯的或混合的）。如果这个过程给每个参与人都只留下一个策略，这个博弈就是占优可解的，并且我们已经找到了一个纳什均衡。更典型地，这个过程仅仅能够缩小策略的范围。在第 5 章中，我们将理性化的策略定义为剔除掉绝非最优反应的纯策略之后剩下的策略集合。现在，我们看到在只有两个参与人的博弈中，我们能够将理性化的策略认为是剔除掉所有被其他纯策略或混合策略占优的策略之后所剩下的策略集合。[8]

7.5 当参与人双方都有三个策略时的混合

正如我们在前边章节 3×2 的策略的例子中所看到的那样，一个在三个纯策略间混合的参与人能够独立地选择其中的两个纯策略的概率（只需它们是非负的，并且它们之和不超过 1）。于是，第三个的概率一定等于 1 减去另两个的概率之和。所以，我们需要两个变量来定义一个混合。[9]当参与人双方都有三个策略时，我们找不到一个不含两个变量代数式的混合策略均衡。在许多情形下，这种代数式仍然是能够处理的。

7.5.1 所有策略的完全混合

考虑足球比赛中罚点球这一简单例子。假定罚球人有三个纯策略：踢向左边、右边或中间，简记为“左”、“中”、“右”。于是，他可以分别用概率 p_L、p_R 与 p_C 在这些策略中进行混合。它们中的任意两个能够被当做独立变量，第三个则由这两个来表示。如果 p_L 与 p_R 被当做两个独立选择变量，于是 $p_C=1-p_L-p_R$。守门员同样有三个纯策略，即移向罚球人的左边（守门员自己的右边）、移向罚球人的右边，或者继续站在中间，简记为“左”、“中”、“右”，并且能够分别以概率 q_L、q_R 与 q_C 在这些策略中进行混合，其中两个概率可以独立选择。

最优反应分析在这部分内容中仍然有用。守门员将选择他的两个独立变量（q_L，q_R）作为应对罚球人的（p_L，p_R）的最优反应，反之亦然。当每个人都选择了最优反应来应对对方的混合策略时，纳什均衡便产生了。然而，由于每个人都是选择两个变量来应对对方的两个变量，我们不能够使用最优反应图，因为这需要四维空间。

取而代之，我们使用对手的无差异准则，它使我们能够聚焦于一个参与人的混合概率。它们应该使得其他参与人混合策略中所有纯策略是无差异的。这就给了我们一组能够解出混合概率的方程。在上述足球例子中，罚球人的（p_L，p_R）将满足两个方程，分别为守门员选择“左”的期望支付等于选择“右”的期望支付，守门员选择“右”的期望支付等于选择“中”的期望支付（于是，选择“左”与选择“中”的期望支付就会自动相等，因此不再是一个独立的方程）。如果有更多的纯策略，需要解得的概率的数量与必须满足的方程的数量也会增加。

表 7—6 以罚球人作为行的参与人，以守门员作为列的参与人表示了这个支付矩阵。由于罚球人希望最大化取胜概率，而守门员则想最小化这个概率，因此这是一个零和博弈。我们从罚球人（行参与人）方面表示支付。例如，如果罚球人踢向左边，而守门员也移向罚球人的左边（左上顶部单元格），我们假定此时罚球人成功的概率为 45%。但是，如果罚球人踢向左边，而守门员移向罚球人的右边或者防守中间，此时罚球人成功的概率为 90%，我们假定有 10%的概率他会踢偏或踢高。如果罚球人踢向右边（矩阵的底部），守门员猜错时他成功的概率为 95%，守门员猜对时他成功的概率为 60%。读者可以列出自己认为更符合实际的不同的支付数字。

表 7—6　　在罚点球博弈中当所有纯策略都被采用时的混合

		守门员		
		左	中	右
罚球者	左	45	90	90
	中	85	0	85
	右	95	95	60

很容易证明博弈没有纯策略均衡。因此假定罚球人以 p_C、p_R 与 $p_C=1-p_L-p_R$ 进行混合。对于守门员的每一个纯策略，这个混合策略获得以下支付：

"左"：$45p_L+85p_C+95p_R=45p_L+85(1-p_L-p_R)+95p_R$

"中"：$90p_L+0p_C+95p_R=90p_L+95p_R$

"右"：$90p_L+85p_C+60p_R=90p_L+85(1-p_L-p_R)+60p_R$

守门员希望这些数字尽可能小。但是，在一个混合策略均衡中，罚球人的混合策略必须使得守门员的纯策略是无差异的。于是，在均衡中，所有的三个表达式必须相等。

令选择"左"与选择"右"的表达式相等并简化，得到 $45p_L=35p_R$，或 $p_R=(9/7)p_L$。接着，令选择"中"与选择"右"的表达式相等并简化，采用刚才得到的 p_L 与 p_R 之间的关系，得到

$$90p_L+95(9p_L/7)=90p_L+85[1-p_L-(9p_L/7)]+60(9p_L/7)$$

或 $[85+120(9/7)]p_L=85$，从而 $p_L=0.355$。

于是，我们得到 $p_R=0.355(9/7)=0.457$，以及 $p_C=1-0.355-0.457=0.188$。罚球人采用这个混合策略应对守门员的任何纯策略以及它们的任意混合所获得的支付能够通过先前的三条支付线的任意一个计算出来，结果是 75.4。

守门员的混合概率能够通过解出罚球人的三个纯策略应对守门员的混合策略的无差异方程来得到。它与解决 5.2 节中相同博弈只在细节上有微小变化。因此，我们省略了细节，只给出答案：$q_L=0.325$，$q_R=0.561$，$q_C=0.113$。罚球人采用任何纯策略应对守门员的均衡混合策略的支付均为 75.4。注意到这个支付与计算罚球人的混合策略时获得的数字相等，这正是零和博弈的最大最小值等于最小最大值的性质。

现在，我们解释这个发现。无论守门员猜对还是猜错，罚球人选择纯策略"右"都会优于选择纯策略"左"（60>45，95>90）。于是，罚球人会以更高的概率选择"右"，相应地，守门员也会以更高的概率选择"右"。然而，罚球人不愿也不会选择纯策略"右"，如果他这样做，守门员也会选择纯策略"右"，这样罚球人的支付只有 60，低于他在均衡中得到的 75.4。

7.5.2 存在一些未使用策略的混合均衡

在前边的均衡中，在混合策略中选择"中"的概率对于每个参与人而言都相当低。（中，中）的组合将导致一个肯定的结果，而罚球人将得到一个相当低的支付，即 0。于是，罚球人选择它的概率将会非常低。而由于守门员关注罚球人更有可能的选择，因此守门员选择中间的概率也会非常低。表 7—6 中博弈的支付数字不会使罚球人避免选择"中"。如果他不选择"中"，守门员也会不选"中"，博弈就会变为一个每个参与人都只有两个基本纯策略（"左"或"右"）的

博弈。然而，在合理的支付组合下，上述结果也可能产生于均衡之中。

我们在表 7—7 中对这个足球博弈做了微小变动。它与表 7—6 中的原始博弈的唯一不同在于将罚球人选择“中”所获得的支付降低了，从 85 降到 70。让我们用 5.1 节中同样的方法计算这个博弈的均衡。这次我们从守门员的角度出发，通过罚球人应该对所有三种纯策略以及它们之间的混合都是无差异的这一条件，试图找到他的混合概率 q_L、q_R 与 q_C。

表 7—7　在罚点球博弈中当不是所有纯策略都被采用时的混合

		守门员		
		左	中	右
罚球者	左	45	90	90
	中	70	0	70
	右	95	95	60

罚球人选择纯策略的支付为：

“左”：$45q_L+90q_C+90q_R=45q_L+90(1-q_L-q_R)+90q_R=45q_L+90(1-q_L)$

“中”：$70q_L+0q_C+70q_R=70q_L+70q_R$

“右”：$95q_L+95q_C+60q_R=95q_L+95(1-q_L-q_R)+60q_R=95(1-q_R)+60q_R$

令选择“左”的式子与选择“右”的式子相等并简化，得 $90-45q_L=95-35q_R$，或 $35q_R=5+45q_L$，令选择“左”的式子与选择“右”的式子相等并简化，得 $90-45q_L=70q_L+70q_R$，或 $115q_L+70q_R=90$。将第一个方程中的 q_R（乘以 2，得到 $70q_R=10+90q_L$）带入第二个方程，得到 $205q_L=80$，或 $q_L=0.390$。将 q_L 这一数值代入任何一个方程可得 $q_R=0.644$。最后，我们用这两个数值得到 $q_C=1-0.390-0.644=-0.034$。由于概率不能为负，很明显存在错误。

为了更好理解在例子中所发生的，首先注意到选择“中”对于罚球人而言是一个比在该博弈的原始版本中更不好的策略，在原始版本中选择“中”的概率已经相当低。但是，让对手各策略无差异（即各方程相等）即意味着罚球人将不得不保持愿意采用这个不好的策略。而只有当守门员采用他的最优反应应对罚球人的策略“中”（守门员也选择“中”）的概率非常低时，这种状况才可能发生。在这个例子中，这种逻辑导致了守门员选择“中”的概率为负。

作为一个纯代数问题，我们推导出的解可能是非常好的，但是它违背了概率理论与真实生活中随机性的要求，即概率必须是非负的。现实中能够做得最好的方法是将守门员选择“中”的概率设得尽可能低，即 0。但是，这会使罚球人也不愿意选择策略“中”。换句话说，我们得到这样一种情况，即每个参与人在混合策略中都有一种纯策略不愿意采用，也就是说，选择它的概率为 0。

每个参与人在他剩余的两种策略（“左”或“右”）间混合是否存在均衡？如果将此看做一个精简了的完整自足的 2×2 博弈，我们就能够轻松地发现它的混合策略均衡。由于读者已经做了许多练习，因此将过程留给读者而只给出结果是可行的。

罚球人的混合概率：$p_L=0.4375$，$p_R=0.5625$；
守门员的混合概率：$q_L=0.3750$，$q_R=0.6250$；
罚球人的期望支付（成功概率）：73.13。

我们只是简单地从直观考虑中剔除两个参与人的策略“中”就得到了这个结果。但是，我们必须检查它是否是上述完整的 3×3 博弈的真正的均衡。也就是说，我们必须确认在给定对方将选择两个策略的混合策略的情况下，两个参与人都不愿意选择他的第三个策略。

当守门员选择这个特定的混合策略时，罚球人从纯策略“中”获得的支付为 0.375×70+0.625×70=70。这个支付低于他从纯策略“左”或“右”以及它们之间的任意混合策略获得的支付 73.13。因此，罚球人不愿意选择策略“中”。当罚球人以前述概率选择“左”或“右”的混合策略时，他应对守门员的纯策略“中”的支付为0.437 5×90+0.562 5×95=92.8。这个数字高于他应对守门员的纯策略“左”或“右”以及它们之间的任意混合策略所获得的支付 73.13，它会使守门员的处境变得更坏。于是，守门员不愿意选择策略“中”。我们所找到的 2×2 博弈的均衡确实就是 3×3 博弈的均衡。

为了能够使得一些策略在均衡中不被采用，我们必须修改或者扩充“对手无差异”原理。参与人无差异于自己所有用过的均衡混合策略，但是对用过和未用过的策略则可以是有差异的。他更喜欢用过的策略而不是未用过的策略。换句话说，为了应对对方的均衡混合策略，他自己的均衡混合策略中用过的所有策略应该给他相同的期望支付，而这个支付应该高于任何未用过的策略所能得到的支付。这称为定理的互补松弛。我们将在第 8 章中考虑它的一般性，在那里这个听起来非常奇怪的术语的内涵将变得更清晰。

7.6 零和博弈中混合策略的证据

从事实验室研究的研究人员通常会轻视混合策略。正如道格拉斯·戴维斯与查尔斯·霍尔特所言：“实验室中，受试者很少考虑到随机因素，并且当事后被告知均衡中包含随机性的事实后，他们常感到惊奇与怀疑。”[10] 当预见的均衡为两个或多个纯策略的混合时，实验的结果确实表明一些群体采用某个纯策略，另一些采用别的纯策略，但这并不构成一个单独参与人的真正的混合策略。当重复进行零和博弈时，单独的参与人会随时间的不同选择不同的纯策略。但是，他们好像是将交替错当成随机了，他们只是经常变换他们的选择，而不是所需要的真正的随机性。

真实的生活仿佛总是领先于实验室，就像下面的故事所描述的一样。[11] 在 20 世纪 40 年代末的马来亚，英国军队需要护卫食品卡车，以免它们遭到游击队的进攻。游击队能够采取大规模的进攻或者制造一次小事故来恐吓卡车司机以使他们不能够继续服役。英军也能采用集中力量护卫或者分散护卫，集中力量护卫有利于反击一次大规模的进攻，分散护卫有利于反击小事故。对于游击队而言，如

果敌人是分散护卫的，一次大规模的进攻将更有效；如果敌人采取集中力量护卫，小事故将更有效。这个零和博弈只有一个混合策略均衡。并不懂博弈论的英军指挥官是这样决策的。每天早晨，当护卫开始时，他抓一片草藏在一只手中，将两只手藏在身后让一个士兵猜哪只手中有草，然后根据士兵是否猜对来决定护卫的形式。由于很难得到精确的支付数字，于是我们不能够说 50—50 是否就是正确的混合，但是这个指挥官已经领会到了真实随机性的需要，并找到了解决它的方法。

支持零和博弈中混合均衡的最好的例子来自体育，特别是职业体育，那些运动员积累了大量的经验。他们对胜利的期望来自获胜后大量的奖金。

马克·沃克（Mark Walker）与约翰·伍德斯（John Wooders）对温布尔登网球赛中的顶级选手的发球与回球进行了检验。[12]他们建立了一个有两个参与人的博弈模型，两个参与人分别为发球者与回球者，他们每人都有两个纯策略。发球者可以选择正手或者反手，回球者能够猜测发球者将选择哪一边并做相应的移动。由于男子单打顶级运动员的发球非常快，回球者不能够在观察到实际的发球方向后再做出反应，于是，回球者必须按照预料的方向进行移动。这个博弈因此是一个同时行动博弈。更进一步，由于回球者希望正确地猜测，而发球者希望误导回球者，这个互动有一个混合策略均衡。

如果在这个发球回球博弈中，网球选手采用均衡混合策略，那么发球者无论选择正手还是选择反手，其赢得这个发球局的概率都是相同的。一场真实的网球比赛包含了在相同的两个选手间进行的上百个博弈，于是有充足的数据来验证前面提到的结论是否成立。马克·沃克与约翰·伍德斯把 10 次比赛发球的结果制成表格。每次比赛包含四种发球回球组合：A 发向 B、B 发向 A、发到场地的左边或者右边（势均力敌的一侧或占优势的一侧）。于是，他们得到 40 种发球情形的数据，并发现其中 39 种，发球人选择正手与选择反手获胜的几率在统计学误差允许的范围内是相等的。

为了能够掌握通常的混合规则以及应对特定选手的正确的混合比例，顶尖选手必须既具备足够的一般比赛经验，又具备应对特定选手的特殊经验。然而，在发球者的选择中有一点与真实的混合策略不同。为了使得发球具备必需的不可预测性，一系列的发球应该不存在任何模式：每次发球的方向应该独立于以前的发球。正如我们在有关混合策略的实践的参考资料中所说的，选手通常会频繁地交替发球，却没有意识到这种频繁地交替与始终坚持一个方向一样都是一种模式。事实上，数据表明网球选手确实会频繁地交替发球。但是，数据同样也说明这种与真实混合策略之间的差别并不足以被对手发觉并加以利用。

在另一项有关运动比赛中的混合策略的研究中，皮埃尔-安德烈·基亚波里（Pierre-André Chiappori）、蒂莫西·格罗斯克洛斯（Timothy Groseclose）与史蒂文·莱维特（Steven Levitt）分析了足球比赛中的点球。[13]他们从欧洲联盟杯比赛中收集了一个大型的数据集，以寻找罚球人与守门员之间的博弈，正如我们在第 5 节中所做的一样。每一个罚球人—守门员组合几乎没有数据可供统计分析采用。因此，他们认为理论的隐含意义针对的是来自不同的罚球人与不同的守门员

的总数据。这项统计工作相当先进，但结果很简单：理论的预言被数据证实。事实上，在他们跟踪的一系列罚球人与守门员的选择中，他们并没有发现频繁地交替。于是，他们的发现表明足球运动员罚点球的行为比网球运动员的发球回球博弈行为更接近于真实的混合策略。

7.7 实践中如何使用混合策略

我们注意到在零和博弈中发现或使用一个混合策略时需要记住一些重要事情。为了在这样一个博弈中有效使用一个混合策略，一个参与人需要做的不仅仅是计算出采用各种行动的均衡百分比。事实上，在我们的网球博弈中，埃弗特并不能够将每次以5/8的概率选择DL、以3/8的概率选择CC机械地转化为五次打底线球、三次打斜线球。为什么不能？因为采用混合策略应该帮助你从让对手的出乎意料的行动中获利。如果你采用一个能够被识别的比赛模式，你的对手就能够发现它并加以利用从而获利。

缺乏固定的模式意味着在任何的选择历史后，下一次选择DL与CC的概率和以前都是一样的。在几次DL偶然成功后，并不意味着CC在下一次成功的机会会更大。在实践中，许多人错误地这样认为，于是他们采用比正确的随机选择所需要的更频繁地交替发球。然而，从观察到的行动中察觉到一个模式是一个需要技巧的统计练习，因为对手在博弈中可能不会这样做。正如我们在第6节中所见到的，沃克与伍德斯在他们对来自于大满贯网球比赛的数据的分析中发现发球人频繁地交替发球，但回球者却不能够发觉到与真实随机性之间的区别。

然而，一个好的策略家并不依靠对手犯错误。为了保证你的混合策略更有效，你需要在博弈的每一次行动中都有一个正确的随机行动模式。例如，你可能希望采用计算机来产生随机数，然后根据这些数来决定你适当的行动选择。如果计算机产生的是在1到100之间的数，而你希望以60∶40的比例混合纯策略A与B，你可以决定对应于1到60之间的任何随机数选择A，对应于61到100之间的任何数选择B。类似地，你可以采用一个像儿童游戏中采用的标色罗盘那样的装置。例如，针对同样的60∶40的混合，你可以在罗盘的圆圈上60%标上蓝色，40%标上红色，即圆圈的前216度标上蓝色，剩下的144度为红色。于是，你可以旋转罗盘的箭头，如果它落在蓝色中间就选A，如果它落在红色中间就选B。手表上的秒针能够起到同样的作用，但重要的是你的手表不能精确到与标准时间同步，以致你的对手能够用一个同样的手表来发现你将要怎样做。

避免一个随机性的可预测系统的重要性在一个零和性质的正在进行的互动行为中是最明显的。由于在这种博弈中参与人利益直接对立，你的对手总是能够从最大限度地发觉你的行动选择中获益。于是，如果你总是在一个规律的基础上采用同样的博弈方式应对对方，他也就总是在寻找破译你随机化行动方式的方法。如果他能够这样做，他就有机会在博弈的未来回合中提高他的收益。混合策略在

单一回合的（有时也称为只有一次的）零和博弈中始终被证明是正确的，因为“出乎意料”的收益是非常重要的。

最后，参与人必须理解并记住这样的事实，即采用混合策略能够防止你的行动被发觉，并给你最大的期望收益，但它仅仅是一个概率的平均。在实际场合中，你可能会获得很差的收益。例如，为了使守门员老实的一码长传球也会在某些特定场合中失败。然而，如果你在一个需要对更高的权威负责的情况下采用一个混合策略，你或许需要事先计划好这种可能性。例如，你可能需要事先向你的教练或老板证明采用这样一个策略是正确的。他们需要理解为什么你会采用混合策略，为什么你期望它能带来最大的期望收益，尽管在某些场合也会带来低收益。甚至这种事先的计划也不能保护你的“声誉”，你应该准备好在坏结果下面对危机。

7.8 总　结

一个参与人宁愿选择与对手相反行动的零和博弈通常都没有纯策略纳什均衡。在这些博弈中，每个参与人希望不能被预测到因而采用一个混合策略，该混合策略是纯策略的概率组合。混合策略是连续策略的一种特例，因此，它相对于连续策略而言有值得独立研究的地方。

最优反应分析能够用来求解混合策略均衡。最优反应图既能够显示一个博弈中的纯策略均衡，又可以用来显示混合策略均衡。混合策略的对手无差异性质说明每个参与人的均衡混合策略对于其他参与人而言，其任何的混合策略都是无差异的。最小最大值分析同样可被用于寻找参与人的均衡混合策略。混合策略均衡解释为在所有参与人都具有有关其他参与人从基本纯行动中进行选择的概率的正确信念时的结果。

当一个或两个参与人具有三个（或更多）策略时，均衡的混合策略也许会对所有的纯策略都有正的概率，也许只包括一个纯策略。如果一个参与人有三个纯策略而其他人只有两个，有三个纯策略的参与人在他的均衡混合策略中通常只用两个。均衡的混合策略也同样是不确定的。

实验室中关于混合策略的证据是模糊的，但在别的现实生活中存在均衡混合策略的证据。网球发球与足球罚点球等体育比赛的例子提供了实际生活的博弈中使用混合策略的证据。并且，在实际生活的博弈中，参与人应该记住一个混合策略的良好应用需要存在随机性的不可预测系统。

关键术语

期望支付（expected payoff）　对手无差异性质（opponent's indifference property）
期望值（expected value）

习　　题

1. “当一个博弈有一个混合策略均衡时，一个参与人的均衡混合策略应该使他应对其他参与人的任意纯策略都能获得相同的期望支付。”这句话正确还是错误？解释理由并给出一个能阐明你的答案的博弈例子。
2. 在本章的图 7—1 与图 7—2 中，你了解了在网球博弈中推导出纳芙拉蒂洛娃最优反应规则的两种不同方式。画一个类似于图 7—1 与图 7—2 的图来表示怎样用这两种不同方式推导埃弗特的最优反应规则。采用代数推导的方式求解其最优反应规则并验证埃弗特的均衡 p 值以及在混合策略均衡中她得到的支付。
3. 找出下面博弈中的混合策略纳什均衡。在每一种情况下，采用第二种方法重新计算每个参与人的均衡混合策略以验证你的答案。

(a)

		列	
		左	右
行	上	4	−1
	下	1	2

(b)

		列	
		左	右
行	上	3	2
	下	1	4

4. 你们中许多人对儿童游戏“剪刀—石头—布”很熟悉。在“剪刀—石头—布”中，两个人同时选择“剪刀”、“石头”、“布”中之一，通常采用手势比画出这三种选择的形状。游戏采用如下记分方式：选择“剪刀”的人击败选择“布”的人（因为剪刀剪碎布），选择“布”的人击败选择“石头”的人（因为布包石头），选择“石头”的人击败选择“剪刀”的人（因为石头会砸坏剪刀）。如果两个人选择了同样的事物，平局。假定赢一次获得 10 分，下面的矩阵表示了博弈的可能结果。

		参与人 2		
		石头	剪刀	布
参与人 1	石头	0	10	−10
	剪刀	−10	0	10
	布	10	−10	0

(a) 假定参与人 2 声称将采用一个混合策略，即以 40%的概率选择“石头”，以 30%的概率选择“剪刀”，以 30%的概率选择“布”。参与人 1 应对参与人 2 的这个策略的最优反应是什么？解释为什么。

(b) 找出这个“剪刀—石头—布”博弈的混合策略均衡。

5. 考虑一个“剪刀—石头—布”的微小变动版本，即参与人 1 具有优势。如果参与人 1 选择“石头”而参与人 2 选择“剪刀”，参与人 1 获得 20 分（超过 10 分）。新的支付矩阵如下：

		参与人 2		
		石头	剪刀	布
参与人 1	石头	0	20	−10
	剪刀	−10	0	10
	布	10	−10	0

(a) 博弈的这个版本的混合策略均衡是什么？

(b) 将你的答案与习题 4 中混合策略均衡的答案比较。你怎样解释均衡策略选择上的不同？

6. 在垒球比赛中，投手与击球手的目标相互冲突。投手希望球能够穿过击球手，但击球手希望击打到球。下面的表格表示了在投手与击球手之间进行的假想的博弈的支付。支付表考虑到了这样一个事实，即击球手击到快球的积分比击到曲线球的积分多。

		投手	
		曲线球	快球
击球手	预料到曲线球	2	−1
	预料到快球	−1	4

(a) 考察这个博弈中是否有纯策略均衡，并找出混合策略均衡。

(b) 画出两个参与人的最优反应曲线，指出在你图表中的混合策略均衡。

(c) 投手能否通过减慢快球的速度（使快球看起来像曲线球）的方式提高在混合策略均衡中的期望支付？假定减慢快球的速度能够将击球手在“预料到快球/快球”单元格中的支付由 4 变为 3。详细解释你是怎样得出答案的，并解释为什么减慢快球的速度能或不能提高博弈中投手的期望支付。

7. 露西（Lucy）提议与查利（Charlie）进行以下博弈：“让我们彼此向对方展示一个便士，每个人选择头像或反面的辫子。如果我们都出头像，我向你付 3 美元；如果我们都出辫子，我向你付 1 美元；如果两人出的不一样，你向我支付 2 美元。”查利的理由如下：“同时出头像的概率为 1/4，这种情况下我得到 3 美元；同时出辫子的概率为 1/4，这种情况下我得到 1 美元；出得不一样

的概率为 1/2，这种情况下我支付 2 美元。因此，这是一个公平的博弈。”他正确吗？如果不正确：（a）为什么不正确？（b）博弈中露西的期望收益是多少？

8. 考虑下面的零和博弈：条目为行参与人的支付，数字 A、B、C 都为正。这些数字中哪些别的关系（例如，$A<B<C$）必须得到满足，如果要使下列情形出现的话？

		列	
		左	右
行	上	0	A
	下	B	C

（a）至少一个参与人有占优策略。

（b）两个参与人都没有占优策略，但存在一个纯策略纳什均衡。

（c）不存在纯策略纳什均衡，但是存在一个混合策略均衡。

（d）假定（c）的条件成立，写出行参与人选择“上”的概率的公式。用 p 表示这个概率，写出它关于 A、B、C 的函数。

9. 假定本章第 5 节中足球运动员罚点球博弈被扩展为罚球者有 6 个策略：分别踢向左上方、左下方、中间上方、中间下方、右上方、右下方，简记为“左上”、“左下”、“中间上”、“中间下”、“右上”、“右下”。守门员仍然有三个策略：移向左方、右方、中间，简记为“左”、“中”、“右”。罚球者成功的概率如下表所示。

		守门员		
		左	中	右
罚球者	左上	0.50	0.85	0.85
	左下	0.40	0.95	0.95
	中间上	0.85	0	0.85
	中间下	0.70	0	0.70
	右上	0.85	0.85	0.50
	右下	0.95	0.95	0.40

这些支付合并了下列信息。即使守门员扑错方向，射高球也会冒射失的风险（于是 0.85<0.95）。如果守门员判断正确，扑住低球的机会高于扑住高球的机会（于是 0.40<0.50）。并且，如果球射向中间而守门员扑向两边，他用脚挡住低球的机会高于挡住高球的机会（于是 0.70<0.85）。

验证一下，在这个博弈的混合策略均衡中，守门员每次以 42.2%的概率选择“左”和“右”，以 15.6%的概率选择“中”，而罚球者每次以 37.8%的概率选择“左下”和“右下”，以 24.4%的概率选择“中间上”。同样证实一

下，均衡中罚球者的支付为 71.8（成功的概率）。（这符合在最近的足球世界杯中，每五个点球中有三或四个成功的事实。）

注意，为了验证上述论断，你所需要做的就是表明给定的数字满足均衡的条件。也就是说，对于对手的均衡混合策略，一个参与人的均衡混合策略中用到的所有策略都使他得到相同的期望支付，它高于那些未用到的策略所带来的支付。

10. 回忆在第 6 章习题 9 中的雷纳与沙格兰之间的决斗博弈。决斗者相隔 10 步，以同样的步伐走向对方，每次 1 步，每一步之后双方都可以开枪射向对方。当一个决斗者开枪时，射中的概率取决于距离，k 步后，射中的概率为 $k/5$。一个决斗者如果被杀死，得到的支付为 -1；如果杀死别人，得到的支付为 1。如果双方都被射死或都未被射死，每人都得到 0。然而，现在假定决斗者的枪装有消音器。如果一个决斗者开枪但没有射中，另一个不知道也就不能采取直到最后一步才开枪以确保射中的策略。每个人在开始时就制定不以别人的中间行动为条件的策略。于是，这是一个同时行动博弈，具有“如果仍然活着，n 步后射击”形式的策略。因此，每个参与人都有五个这样的策略。

(a) 这个博弈的 5×5 的支付表表示如下：

		沙格兰				
		1	2	3	4	5
雷纳	1	0	−0.12	−0.28	−0.44	−0.6
	2	0.12	0	0.04	−0.08	−0.2
	3	0.28	−0.04	0	0.28	0.2
	4	0.44	0.08	−0.28	0	0.6
	5	0.6	0.2	−0.2	−0.6	0

(b) 验证一下，对每个参与人存在一个具有如下策略的混合策略纳什均衡：策略 1 和策略 4 不被采用，策略 2、策略 3、策略 5 分别以 5/11、5/11、1/11 的概率采用。（也就是说，检查这些数字是否满足在一个零和博弈中混合策略纳什均衡的条件。）

(c) 在均衡中，每个参与人的期望支付是多少？

【注释】

[1] 当一个机会事件有两种可能结果时，人们通常会考虑选择某个结果的似然比。如果用 A 与 B 表示两种可能的结果，发生 A 的概率为 p，发生 B 的概率为 $(1-p)$，于是比率 $p/(1-p)$ 就是人们选择结果 A 的似然比，反过来，比率 $(1-p)/p$ 就是人们选择结果 B 的似然比。于是，当埃弗特选择 CC 的概率为 0.25 时，那么她选择结果 CC 的似然比为 1 比 3，不选择结果 CC 的似然比为 3 比 1。这个术语通常用在打赌中，你们中因此而荒废了青春岁月者可能会对此更加熟悉。然而，这个用法不能扩展到有三个或更多可能结果的情形，因此我们在这儿避免使用它。

[2] 博弈理论假定当存在随机的混合策略或结果时，参与人将计算并试图最大化他们

的期望支付。在本章的附录中我们将对此进行深入的讨论，但在这里仍然继续使用它并把它当成重要的前提。“期望支付”中的“期望”这个词语是一个来自概率统计的术语。它仅仅表明一种概率加权平均，并不表示这个支付就是参与人一定或理应能够得到的支付。

[3] 在你试图解决的数字问题中，如果纯策略的两条期望支付线不相交，它表明有一个纯策略针对对手的所有混合策略都是最优的。于是，这个参与人的最优反应就总是这个纯策略。

[4] 有时这种方法会产生错误的结果，例如得出负的或大于 1 的概率。这表明所分析的博弈存在纯策略均衡，你应该回到原来的基本的支付矩阵把它找出来。

[5] 即使当一个参与人只有两个纯策略时，他也许在均衡中不采用某一个。其他的参与人通常会发现一个更好的策略应对第一个参与人所采用的策略。换句话说，均衡的“混合”会变为纯策略的特殊情况。但是，当一个或两个参与人有三个或更多策略时，我们能够有一个真正的混合策略均衡，此时一些纯策略就不起作用。

[6] 我们构造了这条平行线以保证占优。另外一些穿过 DL 与 CC 的交点的直线同样能够占优于 Lob，只需它们的斜率与 Lob 线的斜率相差不大。

[7] 在 5.4 节（表 5—3）中，我们看见列的策略 C4 不是一个最优反应，但它不能被任何一个纯策略 C1、C2 或 C3 占优。现在，我们知道能够通过一个混合策略寻找到占优。在那个博弈中，非常容易地知道 C4 会被 C1 与 C3 的等概率的混合占优。

[8]“绝非最优反应”与“被一个混合策略占优”两者之间的相等在两个参与人的博弈中是毫无问题的，但在多个参与人的博弈中就变得很复杂。考虑有三个参与人 A、B 与 C 的博弈。A 的策略之一（即 A1）也许对于 B 与 C 的任何纯策略或独立混合策略都是绝非最优反应，但是 A1 可能不能够被 A 的任何纯策略或混合策略占优。然而，如果 A1 对于 B 与 C 的任何纯策略或绝对相关混合策略都是绝非最优反应，它一定能够被 A 的一个纯策略或混合策略占优。一个完整的处理过程需要更高级的博弈理论，因此我们仅仅是提及它。可参考 Andreu Mas-Colell，Michael Whinston，and Jerry R. Green，*Microeconomic Theory*，New York：Oxford University Press，1995，pp. 242 - 245 and 262 - 263。

[9] 更一般地，如果一个参与人有 N 个纯策略，那么他的混合策略有（$N-1$）个独立变量，或者称为“选择的自由度”。

[10] Douglas Davis and Charles Holt，*Experimental Economics*，Princeton：Princeton University press，1993，p. 99.

[11] R. S. Beresford and M. H. Peston，“A Mixed Strategy in Action”，*Operations Research*，vol. 6，no 4，December 1995，pp. 173 - 176.

[12] Mark Walker and John Wooders，“Minimax Play at Wimbledon，” *American Economic Review*，vol，91，no. 5，December 2001，pp. 1521 - 1538.

[13] Pierre-andré Chiappori，Timothy Groseclose and Steven Levitt，“Testing Mixed Strategy Equilibria When Players Are Heterogeneous：The Case of Penalty Kicks in Soccer，” *American Economic Review*，vol. 92，no. 4，September 2002，pp. 1138 - 1151.

附录　概率与期望效用

要计算本章的期望支付及混合策略均衡，我们必须做一些简单的概率运算，有一些概率计算的简单公式，你大概早已熟悉，但我们还是在此做一些概略的叙述及解释，当做复习，我们也会叙述如何计算随机变量的期望值。

我们也会讨论以期望效用方法计算期望支付。当某博弈行动的结果不明确时，可能因为你的对手使用混合策略，或是本质上不确定的因素，你不想让自己的货币期望支付最大化，反而注重支付的风险性。我们在第 2 章提到过，可以用适当的金钱支付的非线性函数期望值来处理此问题，我们在此提供简单的讨论及做法。

本附录浅显易懂，但要勤加练习才能真正学会，接下来的章节，特别是第 8 章、第 9 章和第 13 章将给你许多练习的机会。

7A.1　概率基本代数

概率给人的基本直觉是：一个事件在一个大概率集合里发生的频率，通常，这个大集合里的任何事件发生的可能性都一样，所以找出我们感兴趣事件的概率很简单，就是用事件包含的样本个数除以总样本数。[1]

在任何标准的 52 张牌中，会有标准的 4 种花色（梅花、方块、红桃、黑桃），每种花色有 13 张（A 到 10 点，以及人头牌——J、Q、K），我们可以问，

从牌中抽出某张特定牌的可能性是多少？例如，抽中黑桃的可能性是多少？抽中黑色牌的可能性是多少？抽中 10 点的可能性是多少？抽中黑桃 Q 的可能性是多少？等等。我们需要了解概率的计算，才能回答这类问题。如果我们有两副牌，一副是绿背的，另一副是蓝背的，我们甚至可以问更复杂的问题（如，从每副牌中各抽出一张，同时是方块 J 的可能性是多少?），同样，我们仍用概率代数回答这类问题。

通常，**概率**（probability）衡量的是一个特定事件或一组特定事件发生的可能性。从一副牌中抽出黑桃的可能性也就是“抽中黑桃”事件发生的概率，在这里集合中有 52 个元素——相同概率事件的总数，“抽中黑桃”事件对应于包含 13 个特定元素的子集合，所以有 52 分之 13 的概率抽中黑桃，也就是 13/52＝1/4＝25％；从另一个角度来看，考虑每种花色有 13 张，所以你抽出某特定花色的概率为四分之一，也就是 25％，当然，如果你抽很多次（都从一副完整的牌中抽取），那么 52 次里面你不会恰恰抽中 13 次黑桃，你会随机地多抽一点或少抽一点，但在大的采样库中，抽中黑桃的平均概率是 25％。[2]

我们用一般的术语来表达概率代数的思想从而推出一些公式，这样你就可以直接套用它们而无须每次都从头做起。我们会围绕从两副牌（例如蓝底牌和绿底牌）中抽牌的问题，组织一次对概率公式的讨论[3]，而且这样的方法也会为你提供一些具体且一般的计算公式。这样，你就可以把抽扑克牌的原理类推到其他的概率问题中。这里有一点要注意：通常我们以百分比表示概率，但代数中需要以小数点表示，所以算数时用 13/52 或 0.25 而不是 25％，两种都可以代表概率，视情况而定，但意思都一样。

7A. 1. 1　加法律

我们的第一个问题是：如果从一副牌中抽一张，抽到黑桃的可能性是多少？抽到不是黑桃的可能性是多少？我们已经知道抽到黑桃的概率是 25％，但抽到不是黑桃的概率是多少呢？由于抽到梅花或方块或红桃跟抽到黑桃的可能性一样大，因此很显然抽不到黑桃的概率比任何构成它的子事件的概率要大，此概率是 13/52(梅花)＋13/52(方块)＋13/52(红桃)＝0.75，这个“或”字暗示概率应该相加，因为我们想知道抽中任何这三种花色的机会。

我们可以用更简单的方法找到答案，请注意，75％的可能性抽不中黑桃，也就是抽中“不是黑桃”的概率是 75％(100％－25％)，或更正式地，1－0.25＝0.75。跟一般概率运算的例子一样，由于不同的思考方式，同样的结果可以由两种不同的路径取得，接下来的例子会清楚地显示不同计算方法的工作量大不相同。当你的经验增加时，你会发现并记住简单的方法或捷径，同时，不同路径导致的答案是一样的，这点毋庸置疑。

归纳上述的计算，当你把你感兴趣的集合 X，分成数个子集合 Y，Z……且这些子集合互相不重叠［用数学术语来说就是**不相交**（disjoint）］时，则所有子集合发生的概率加起来一定等于母集合发生的概率；如果母集合包括所有可能的

结果，则它发生的概率是 1。换句话说，如果集合 X 包含若干个不相交的子集合 Y，Z……则 X 发生的概率是这些子集合发生的概率的总和。用 Prob(X) 表示 X 发生的概率，则**加法律**（addition rule）可表示成 Prob(X)＝Prob(Y)＋Prob(Z)＋……

练习：用加法律找出以下概率：从两副牌中各抽出一张牌，这两张牌有相同的人头。

7A.1.2 修正加法律

上一节的分析只涵盖母集合及不相互重叠的子集合，但如果我们问，从一副牌中抽出一张，是黑桃或 A 的可能性是多大？这个“或”字同之前一样，暗示我们概率相加，但这个例子中，“黑桃”与“A”有重复的地方，也就是“黑桃”和“A”同时存在两个集合里，所以“黑桃”与“A”是相交的子集合。如果我们把抽到黑桃（13/52）及抽到 A(4/52)的概率加起来，我们得到 17/52，也就是说有 17 种黑桃或 A 的可能，但事实上只有 16 种，13 张黑桃（包括 A）及另外 3 张不同花色的 A，答案 17/52 错误是因为重复计算了黑桃 A 两次。为了得到正确的概率，我们必须把两个子集合重叠的部分减掉，所以抽中黑桃或 A 的概率就是，抽中黑桃的概率加上抽中 A 的概率减掉重复的概率，所以是 13/52＋4/52－1/52＝16/52＝0.31。

归纳起来说，如果你把你感兴趣的集合 X 分成几个子集合 Y，Z，……这些子集合可能重叠，则每个集合概率的加总减掉重叠部分，就是集合 X 整个的概率。更正式的说法，**修正加法律**（modified addition rule）阐述为：如果集合 X 包含了如果任何两个相交的子集合 Y 及 Z，则 X 的概率是个别子集合概率的总和减掉 Y 与 Z 同时发生的概率：Prob(X)＝Prob(Y)＋Prob(Z)－Prob(Y 与 Z)。

练习：用修正加法律找出以下概率：从两副牌中各抽出一张牌，抽中至少一张人头的概率。

7A.1.3 乘法律

现在问，从两副牌中各抽出一张牌，都是黑桃的可能性有多大？如果第一副牌和第二副牌都抽出黑桃，那么该事件就发生了。这里是“和”而不是“或”，暗示了应该用乘法而不是加法，所以从两副牌中都抽出黑桃的概率是每副牌抽出黑桃概率的乘积，或是(13/52)×(13/52)＝1/16＝0.062 5 或 6.25%。对于抽两张黑桃比抽一张黑桃的概率还低，我们并不惊讶（用你对结果的直觉再确认一次）。

就如同加法律要求所有事件不相交一样，乘法律也要求它们互相独立。如果我们把一事件分成几个子集合 Y，Z……如果某一集合发生的概率不影响另一集合发生的概率，则它们是互相独立的集合。我们的例子——第一副牌和第二副牌

都抽出黑桃——满足了独立的情况，也就是从第一副牌抽出黑桃不影响从第二副牌抽出黑桃的概率。如果我们从同一副牌中抽出两张黑桃，则一旦我们抽出第一张黑桃（概率是 13/52），第二张抽出黑桃的概率就不再是 13/52 了（事实上是 12/51），所以从同一副牌中抽出两张黑桃不是**独立的事件**（independent events）。

对**乘法律**（multiplication rule）的正式说明告诉我们：如果事件 X 是几个同时发生的独立事件 Y，Z……则 X 的概率是 Y，Z……概率的乘积：$\text{Prob}(X)=\text{Prob}(Y)\times\text{Prob}(Z)\times\cdots\cdots$

练习：用乘法律找出以下概率：从两副牌中各抽出一张牌，从第一副牌中抽出红花色牌，从第二副牌中抽出人头牌。

7A.1.4 修正乘法律

如果我们问的是某一事件发生的概率，而该事件由两个非独立事件组成，例如我们问，我们抽一张牌，是黑桃又是 A 的可能性为多大？思考片刻之后，我们知道该事件发生的概率即为抽中黑桃而且又是 A 的概率。抽中黑桃的概率是 13/52=1/4，黑桃中抽中 A 的概率是 1/13，“又是”两个字暗示我们要取两概率的乘积：(13/52)×(1/13)=1/52。

当然，如果我们用以下问法答案也会一样：抽中黑桃 A 的概率是多少？该概率的计算直截了当，52 张牌中只有 1 张牌是黑桃 A。你会发现，你对问题的看法会影响你解决问题的方法。

用概率的术语来说，假设在另一事件已发生的情况下（抽中黑桃），某一特定事件发生的概率（抽中 A）叫做**条件概率**（conditional probability），也就是在抽中黑桃的条件下抽中 A 的概率。正式地说，**修正乘法律**（modified multiplication rule）告诉我们，如果 X 事件包含 Y 跟 Z，则 X 发生的概率等于以下两者的乘积：(1) 单独 Y 发生的概率 (2) 在 Y 发生的条件下 Z 发生的概率，或称在 Y 发生的情况下，Z 发生的条件概率，即 $\text{Prob}(X)=\text{Prob}(Z)\times$ Prob(在 Z 发生条件下的 Y)。

我们可以用第三种方式描述本例的概率，抽中 A 的概率是 4/52，而这 A 花色是黑桃的条件概率是 1/4，所以抽中黑桃 A 的概率是(4/52)×1/4。我们用更一般的方式来表达刚才介绍的方法：$\text{Prob}(X)=\text{Prob}(Z)\times$ Prob(在 Z 发生条件下的 Y)。

练习：用修正乘法律找出以下概率：从一副牌中抽出两张牌，第二张牌是红桃 J。

7A.1.5 结合律

现在我们提出目前为止最困难的问题，需要同时用到加法律（或修正加法

律）及乘法律（或修正乘法律）。这一问题是，从两副牌中各抽出一张，抽中至少一张黑桃的概率是多少？跟平常一样，我们可以从几个角度来看这个问题。首先考虑所有的可能性，有三种：从第一副牌抽出黑桃而没有从第二副牌抽出黑桃（简称“黑桃、无”）；没从第一副牌抽出黑桃而从第二副牌抽出黑桃（简称“无、黑桃”）；从每副牌都抽出黑桃（简称“黑桃、黑桃”）。这个事件包含这三种概率，而每一种又分别含有两个独立事件，很明确地要用“或”及“和”作为计算的导引。我们找出三种可能性的概率（每一种都包含两个概率的乘积），把它们加起来：

$$1/4\times3/4+3/4\times1/4+1/4\times1/4=7/16=43.75\%$$

使用第二种方法，我们要了解“至少一张黑桃”跟“没有任何黑桃”是分离的事件（不互交事件），两个合起来代表确定事件，所以“至少一张黑桃”的概率是1减去“没有任何黑桃”的概率，而“没有任何黑桃”发生的概率是第一副牌没抽中黑桃而第二副牌也没抽中黑桃的概率，也就是 $3/4\times3/4=9/16$，所以“至少一张黑桃”的概率是 $1-9/16=7/16$，答案跟前面一样。

最后，我们正式地阐述概率的**结合律**（combination rule）：如果事件 X 包含几个分离的事件（不互交事件）Y，Z……而事件 Y 是几个同时发生的事件 Y_1，Y_2……事件 Z 是几个同时发生的事件 Z_1，Z_2……如此等等，则 X 的概率是 Y，Z……概率的和，而 Y，Z……的概率又分别是 Y_1，Y_2……的积及 Z_1，Z_2……的积，即

$$\begin{aligned}\mathrm{Prob}(X)&=\mathrm{Prob}(Y)+\mathrm{Prob}(Z)+\cdots\cdots\\&=\mathrm{Prob}(Y_1)\times\mathrm{Prob}(Y_2)\times\cdots\cdots+\mathrm{Prob}(Z_1)\times\mathrm{Prob}(Z_2)\times\cdots\cdots\end{aligned}$$

练习：假设我们现在有三副牌，找出以下概率：从三副牌中各抽一张，抽中至少一张黑桃。

7A.1.6 期望值

如果一个数字变量会发生改变（例如赢的钱数或降雨量），而它的值有 n 种可能 X_1，X_2，…，X_n，概率分别是 p_1，p_2，…，p_n，则期望值的定义是所有可能值经过概率加权后的平均，也就是

$$p_1X_1+p_2X_2+\cdots p_nX_n$$

例如，在掷两枚硬币时，你打赌，如果掷出两个头你赢5美元，一头一尾你赢1美元，都是尾则没有钱。用之前学过的概率运算法则，可知这三个事件的概率分别是0.25、0.50、0.25，所以你的期望值是

$$0.25\times5+0.50\times1+0.25\times0=1.75$$

在博弈理论里，这一数值叫做支付，可以表示大小、金钱或效用。随后我们将讨论效用。为了配合期望值的意义，我们将它称作期望支付或期望效用（expected utilities）。

7A.2 对风险的态度及期望效用

第2章曾指出在博弈中用概率计算期望值的困难性。现在考虑一个游戏，参赛者会赢钱或输钱，我们简单地以钱的多寡来表示支付。如果某一参赛者有75%的概率得不到钱而25%的概率得到100美元，则期望支付是概率加权平均：期望值是不同支付的不同概率的加权平均。在这个例子，我们有75%的概率得到0美元，收益为$0.75\times0=0$，加上25%的概率得到100美元，收益为$0.25\times100=25$，这跟保证会得到25美元的博弈的支付一样。对这两种方法（同样支付但不同风险）保持中立的人，称作**风险中性**（risk-neutral）。在我们的例子，一种选择是规避风险（选择确定的25美元），另一种选择是冒险，有0.75的概率一无所获，有0.25的概率获得100美元。相反，**风险规避**（risk-averse）的人——面对两个可获得同样期望支付的选择，他会选择比较没有风险的那个，在我们的例子，宁愿获得稳当的25美元，也不愿面对100美元或0美元的风险。这样的厌恶风险的行为很普遍，而我们应该有一套考虑风险的决策理论。

我们曾在第2章说过，简单修正一下计算支付的方式就可以让我们绕过这一困难。我们可以不用金钱总数来衡量支付而用金钱数量的非线性函数，这里将显示如何构造函数及为什么它可以解决问题。

假设当一个人得到D美元，我们定义支付不是D，而是$\sqrt{D}$，所以0美元的支付是0，100美元的支付是10。这一转换并不改变一个人对0美元与100美元的态度，只是对支付做特定缩放罢了。

现在考虑有风险的选择，也就是得到100美元的概率是0.25，得到0美元的概率是0.75，缩放之后，期望支付是$0.75\times0+0.25\times10=2.5$。这一期望支付跟某人得到6.25美元的意义一样，因为$\sqrt{6.25}=2.5$，一个确定会得到6.25美元的人也会得到2.5的支付。换句话说，平方根支付函数让获得6.25美元的效果，跟以0.25概率获得100美元的效果一样。这相同的效果反映了强烈的风险规避意愿，也就是某人愿意放弃25美元跟6.25美元的差异以避免风险，图7A—1显示了此非线性函数（平方根）。

如果我们用三次方根来做非线性函数呢？则100美元的支付是4.64，博弈的期望支付是$0.75\times0+0.25\times4.64=1.16$，也就是1.56的立方根，所以某人会选择稳当的1.56美元而不选择平均的25美元，这样的人是极度风险规避的（比较立方根的图与平方根的图）。

而如果用x^2来做非线性函数呢？博弈的期望支付是$0.75\times0+0.25\times10\,000=2\,500$，也就是50的平方。所以某人在稳当地得到50美元和博弈平均得到25美元之间保持中立，这人一定是风险偏好者，因为他不放弃冒风险能带来的金钱，相反，如果不冒风险，他就能得到额外25美元的补偿，图7A—2显示此非线性函数。

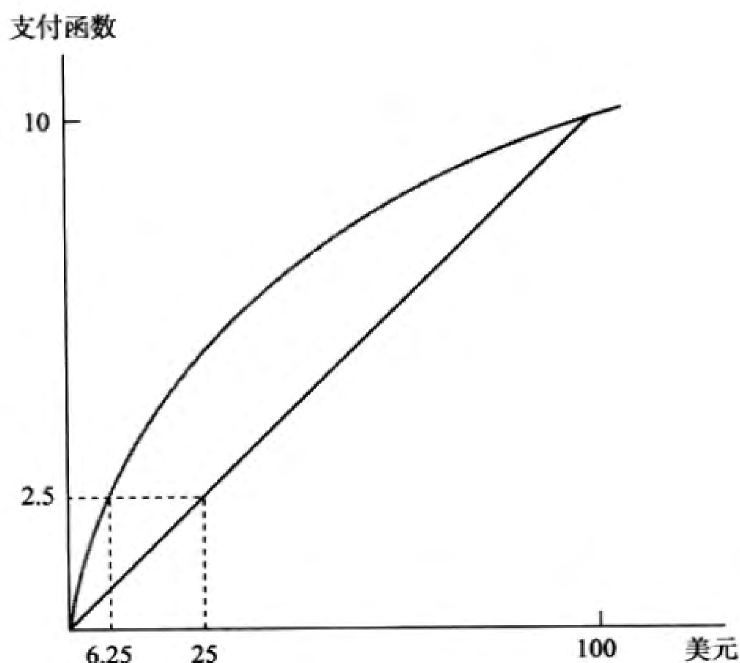


图 7A—1　凹函数：风险规避

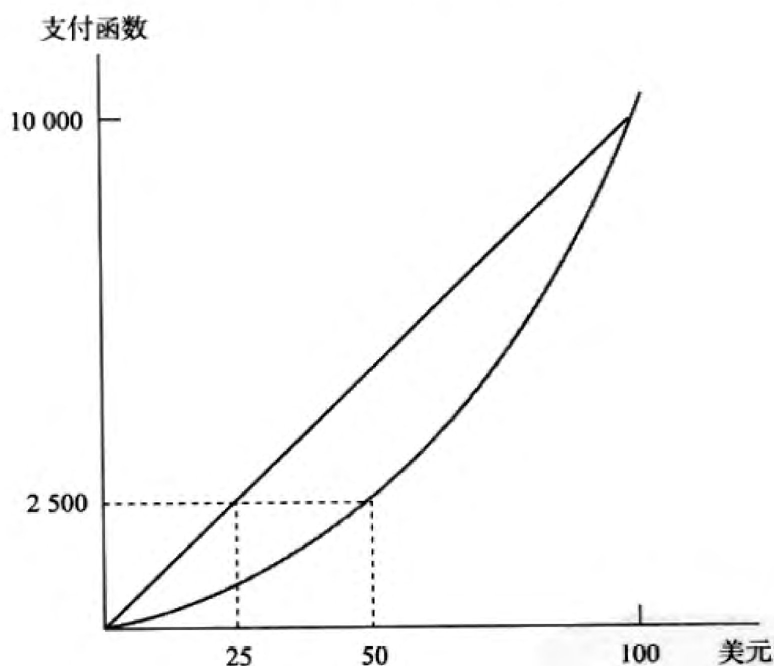


图 7A—2　凸函数：风险偏好

所以借由不同的非线性函数，我们可以捕捉不同的厌恶风险或偏好风险行为，如图 7A—1 的凹曲线是风险规避，7A—2 的外凸曲线是偏好风险，你可以用不同的非线性函数来做实验——例如对数、自然指数、n 次方根或级数——看看它们能暗示关于风险态度的什么信息。[4]

这种评估风险的方法在决策理论上有悠久传统，称作期望效用方法；用非线性函数方式把支付表示成金钱数量的函数，称作**效用函数**（utility function）。平方根、立方根、平方等都是例子，而效用值的概率加权平均称作期望效用。高的

期望效用比低的期望效用好。

几乎所有博弈理论都是基于期望效用方法的，它虽然不是完美无缺，但的确管用。我们就介绍到此，不做进一步讨论。[5]然而，我们会在第 8 章以简单的例子说明期望效用方法的缺陷。

7A.3 总　　结

一个事件的概率是指在一个大概率集合里它发生的机会。概率可以按照一些规则相加，加法律是指某些不相交事件可能发生的概率，也就是这些事件概率的相加，修正加法律推广到了相交的事件。修正乘法律推广了乘法律，采用条件概率，使它能够应用于缺乏独立性的情况。

采用期望金钱支付来判定结果假设了风险中性行为。采用期望效用的方法可以研究风险规避的情况，它需要采用一个风险效用函数，这个函数是金钱支付的凹函数，并用其概率加权平均来度量期望支付。

关键术语

加法律（addition rule）
结合律（combination rule）
条件概率（conditional probability）
不相交（disjoint）
期望效用（expected utility）
独立事件（independent events）
修正加法律（modified addition rule）
修正乘法律（modified multiplication rule）
乘法律（multiplication rule）
概率（probability）
风险规避（risk-averse）
风险中性（risk-neutral）
效用函数（utility function）

【注释】

[1] 当我们说“随机”时，是指结果不能被系统化地侦测，或是不能用科学方法预测及计算。事实上，骰子及铜板的运动完全由物理法则决定，技术好的人可以操纵纸牌，但一般来说，掷铜板、掷骰子或是洗牌都是可以产生随机结果的方法，但随机化没有你想象的那样容易达成，例如一副完整的牌，一分为二，然后依次序从每半副中拿出一张，组成新牌，直到每半副都被拿完，这似乎是打乱原来牌序的好方法，但康乃尔大学的数学家 Persi Diaconis 证明，经过八次洗牌后，原来的牌序会被重建。对一些不完整的牌，他发现洗牌六次还有某些次序，但第七次却变成没有次序，请参考“How to Win at Poker, and other Science Lessons,” *Economist*, October 12, 1996。关于这类议题的有趣讨论，请参考 Deborah J. Bennett, *Randomness*, Cambridge: Harvard University Press, 1998, chaps. 6—9。

[2] Bennett, *Randomness*, chap. 4 and 5 提供了计算此类概率问题的几个例子。

[3] 想获得关于加法律及乘法律的更多详细的说明，多练习这些规则的应用，我们建

议参考 David Freeman, Robert Pisani, Rogert Purves, *Statistics*, 3rd ed., New York: Norton, 1998, chap. 13 and 14。

[4] 期望效用及风险态度的额外资讯可以在许多微观经济学课本找到，例如 Hal Varian, *Intermediate Microeconomics* 5th ed., New York: Norton, 1987, pp. 216 - 220; Walter Nicholson, *Microeconomic Theory*, 7th ed., New York: Dryden Press, 1988, pp. 211 - 226。

[5] 请参考 R. Duncan Luce and Howard Raiffa, *Games and Decisions*, New York: Wiley, 1957, chap. 2 and app. 1, for an exposition; Mark Machina, "Choice Under Uncertainty: Problems Solved and Unsolved," *Journal of Economic Perspective*, vol. 1, no. 1, Summer 1987, pp. 121 - 154。前书提供解说，后文则提供评论与替代方法。虽然决策理论有了长足的进步，但是尚未对博弈理论产生显著影响。

第 8 章　采用混合策略的同时行动博弈（Ⅱ）：非零和博弈、（Ⅲ）：一般性讨论

在第 7 章中，我们考虑了参与人在均衡中采用混合策略的零和博弈。在那些博弈中内在的直接冲突使得随机性行动对参与人具有吸引力。并且在那一章中，我们对零和博弈的分析说明了混合策略在博弈中是如何使对方保持猜疑来帮助参与人的。

我们同样在 7.3 节中指出混合策略也能够在非零和博弈中出现。在非零和博弈中，参与人在利益上没有明显的冲突并且不会因为寻常的理由对别人隐藏自己的兴趣。结果，在非零和博弈中不存在使别的参与人保持猜疑的寻常的要求。同时行动仍然能使参与人对对手的行动具有不确定的信念，因此也不确定自己的最优行动。当参与人获得主观不确定但正确的关于行动的信念时，非零和博弈中就会出现混合策略均衡。在本章中，我们将仔细研究这种类型的混合策略均衡。

本章也会涉及与混合策略均衡相关的更深层的内容。我们特别考虑了一些具有这种均衡的博弈中出现的重要的违反直觉的结果。我们也提供一节讨论所有参与人有几个纯策略的混合策略的一般性理论。这些补充进来的内容能帮助你更直观地理解混合策略均衡，并且锻炼你对具有这种均衡的博弈的分析技巧。

8.1 不确定信念支撑的混合策略

在同时行动非零和博弈中，当每个参与人选择自己的策略时，由于对别人的选择难以确定从而也不确定自己的选择，就会出现混合策略。例如，在 4.7 节中的约会博弈中，哈里不确定萨莉将会选择哪个咖啡店。这种不确定会使他不知如何选择。那么萨莉也会由于同样的原因而不确定哈里将要去的地方。于是，博弈会具有一个混合策略均衡，其中这种相互的不确定恰好处于合理的水平上。也就是说，参与人彼此都对对方的不确定行动具有正确的信念。

8.1.1 哈里会碰上萨莉吗?

我们采用约会博弈的安全版本来演示这种类型的均衡。为了方便你的阅读，我们将表 4—12 复制为下面的表 8—1。我们首先从萨莉方面来考虑这个博弈。如果她确信哈里将要去星巴克，她也同样会去星巴克。如果她确信哈里将要去本地咖啡店，她也同样会去。但是如果她不能确定哈里的选择，她自己的最优选择又将会是什么呢?

表 8—1 安全版本

		萨莉	
		星巴克	本地咖啡店
哈里	星巴克	1，1	0，0
	本地咖啡店	0，0	2，2

为了回答这个问题，我们必须给萨莉意识中的不确定一个更精确的含义。(在概率统计理论中，这个不确定的技术术语就是她的主观不确定 [subjective uncertainty]。) 我们通过规定萨莉认为哈里将选择某家或别的咖啡店的概率来赋予不确定更精确的含义。哈里选择本地咖啡店的概率为 0 到 1 之间的实数（也就是说，在 0%到 100%之间)。我们用代数来覆盖所有的可能情况。令 p 表示萨莉心目中哈里选择星巴克的概率，变量 p 为在 0 到 1 之间的任何实数，则 $(1-p)$ 表示萨莉心目中哈里选择本地咖啡店的概率。换句话说，我们这样描述萨莉的主观不确定：她认为哈里将采用混合策略，分别以 p 和 $(1-p)$ 的概率选择纯策略星巴克和本地咖啡店。我们将这个混合策略称为哈里的 p—mix，尽管目前它仅仅是萨莉心目中的一个想法。

假定萨莉正是按照这种方式思考哈里的选择的，她自己的最优行动又是什么? 为了找到答案，我们必须评估当萨莉面对哈里的 p—mix 时，两种可能的选择带来的支付。这里，我们考虑表 8—1 中博弈矩阵的一个改进版本，这次只表

示出萨莉的支付并增加了一行哈里的p—mix，新矩阵如表 8—2 所示。

表 8—2　　萨莉对哈里的 p—mix 的反应

		萨莉	
		星巴克	本地咖啡店
哈里	星巴克	1	0
	本地咖啡店	0	2
	p—mix	p	2 $(1-p)$

给定萨莉关于哈里的选择有如下思考，她现在想："如果我选择星巴克，那么哈里将以 p 的概率在那儿，则我也会以 p 的概率得到支付 1；而哈里将以 $(1-p)$ 的概率在本地咖啡店，则我也会以 $(1-p)$ 的概率得到支付 0。"期望值为 $p\times1+(1-p)\times0=p$，这就是我们在表 8—2 中对应于哈里的 p—mix 与萨莉选择星巴克的单元格中所记录的。同样，她选择本地咖啡店应对哈里的 p—mix 的期望支付为 $p\times0+(1-p)\times2=2(1-p)$，这个数值也表示在了表 8—2 相应的单元格中。

现在，我们可以找到像第 7 章中那样的萨莉的最优反应。图 8—1 的左边一幅图表示了萨莉采用两个纯策略应对哈里的 p—mix 的期望支付。上升的为选择星巴克获得的支付直线 p，下降的为选择本地咖啡店获得的支付直线 $2(1-p)$，两条直线相交于 $p=2/3$ 处。在交点的左边，萨莉选择本地咖啡店会得到更高的期望支付；在右边，萨莉选择星巴克会得到更高的期望支付；在 $p=2/3$ 处，萨莉从两种纯策略中获得的期望支付相同，因此她的 q—mix 策略中的任何混合效果都是相同的，这里 q 为她选择星巴克的概率。

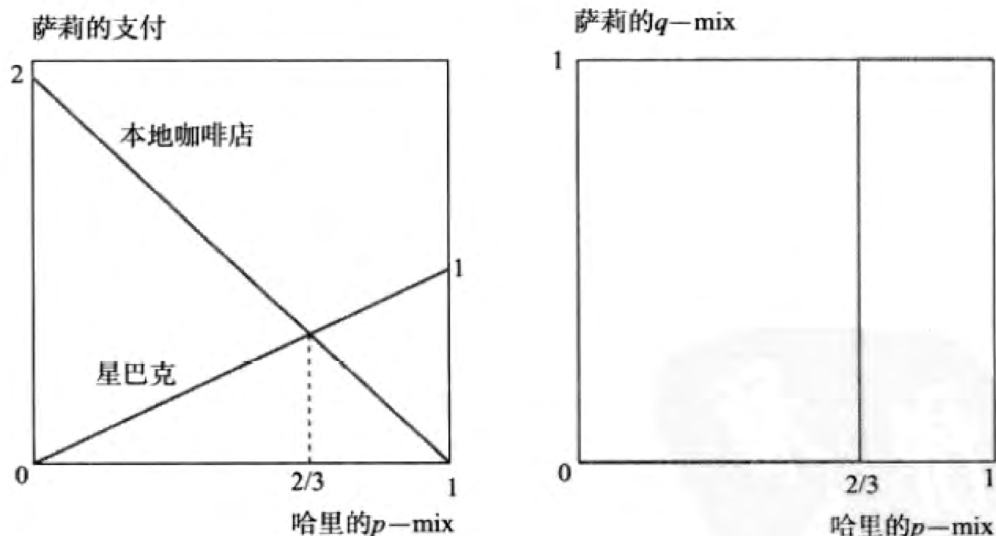


图 8—1　萨莉的最优反应

图 8—1 的右边一幅图将萨莉针对不同的 p 值获得的期望支付转化为最优反应曲线。当 $p<2/3$ 时，萨莉的最优反应为纯策略本地咖啡店（$q=0$）；当 $p>2/3$时，萨莉的最优反应为纯策略星巴克（$q=1$）；当 $p=2/3$ 时，所有的 q 值对于萨莉而言都是一样的，因此她的最优反应为星巴克与本地咖啡店的任意组合，

即在 $p=2/3$ 处的这一部分最优反应曲线是以从 $q=0$ 到 $q=1$ 处的垂直线段所表示的。

关于萨莉的最优反应曲线的直观解释很简单。数值 p 代表萨莉所估计的哈里去星巴克的概率。如果 p 较高，代表哈里去星巴克的可能性更大，于是萨莉也将去那里；如果 p 较低，萨莉将去本地咖啡店。在平衡点或叫无亏损点，p 为 2/3 而不是 1/2，这是因为萨莉在本地咖啡店遇到哈里的支付高于在星巴克遇到哈里的支付。因此，对于萨莉的决策而言，为了抵消支付间的不平衡，哈里会去星巴克的概率应超过 1/2。

萨莉的最优反应曲线的垂直部分的存在具有重要的含义。如果$p=2/3$，所有的纯策略与混合策略对于萨莉而言都是一样的，因此她可能会在两个行动中不知所措或在它们之间随机选择。萨莉意识中关于哈里的行动的主观不确定能够导致一个在她行动上的真实的或客观的不确定，而哈里也是一样，这就是他们的主观不确定如何结合以支撑一个混合策略均衡的。在这个博弈中的一个混合策略纳什均衡将是这样一种情形，即每个参与人都有关于对方行动的恰当的不确定性以支撑自己的随机性。

正如我们在 7.2.1 节中所做的一样，我们通过计算哈里的最优反应并重叠两条曲线来发现混合策略均衡。由于两个参与人的支付是对称的，哈里的最优反应曲线看起来就像萨莉的最优反应曲线颠倒了坐标轴。图 8—2 同时表示了两条最优反应曲线。萨莉的最优反应曲线为黑色的粗线，哈里的反应曲线为灰色的粗线。

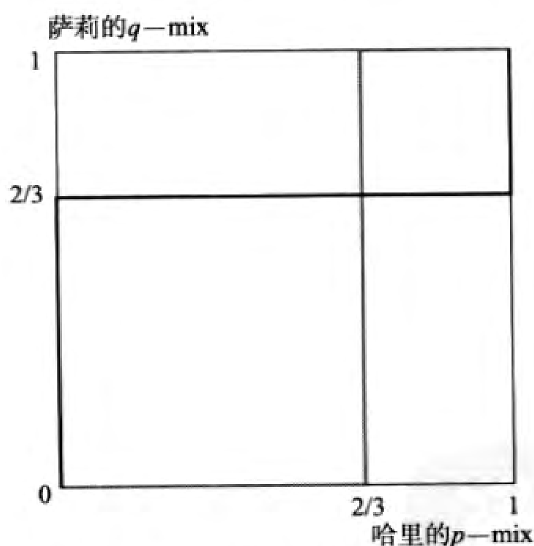


图 8—2　安全博弈中的混合策略均衡

两条最优反应曲线相交于三点。第一点为右顶端，此时 $p=1$，$q=1$。这一点对应于每个参与人都确定选择星巴克，并在主观信念中认为别人也会这样做。这是事物的自维持状态。第二点为左底部，此时 $p=0$，$q=0$。这一点不将任何概率放在去星巴克上，也就是说，每个参与人都确定会去本地咖啡店，并在主观信念中认为别人也会这样做。这同样是事物的自维持状态。它们正是我们在第 4 章中找到的这个博弈的两个纯策略纳什均衡。

但是在图 8—2 中存在第三个交点，此时 $p=2/3$，$q=2/3$。这个交点确定了我们在第 4 章分析中未曾发现的博弈的第三个纳什均衡。在这个均衡中，两个参与人都采用混合策略。每个人都不能确定对方行动的选择，并存在一个他们主观不确定的均衡上的平衡。由于萨莉认为哈里将以 2/3 概率去星巴克在两个咖啡店之间混合，她也最好是以 2/3 的概率混合，并且哈里也同样如此，于是他们的信念在均衡中被证实。

记得我们对纳什均衡的定义并不要求一个参与人的均衡策略严格优于他的其他策略，而仅仅是要求没有其他策略优于均衡策略。这里，给定萨莉关于哈里混合策略的信念，萨莉的所有策略，无论纯策略还是混合策略，对于她都是一样的。她没有明确的理由选择以精确的比例 2：1 喜欢星巴克。然而，她这样做足可支撑一个均衡，其他任何选择都将不是一个均衡。看起来仿佛在支持均衡上会出现一个微小的争论，但是如果你曾经因为不知如何选择而不知所措，你就应该理解在实践中这样一个均衡是如何产生的。

这个均衡有一个重要的缺陷。如果双方都各自以 $p=2/3$ 与 $q=2/3$ 进行随机选择。那么他们只有在双方都恰巧选择了同一咖啡店时才会相遇。因此，他们在星巴克相遇的概率仅为 $pq=4/9$，他们在本地咖啡店相遇的概率为 $(1-p)(1-q)=1/9$。如果他们做不同的选择，将不会相遇。这种不幸的概率为 $1-4/9-1/9=4/9$。于是，每个人在混合策略均衡中的期望支付为 $(4/9\times1)+(1/9\times2)+(4/9\times0)=6/9=2/3$。两个纯策略纳什均衡将会给每个参与人更好的支付，甚至这两个均衡中较差的一个（$p=1$，$q=1$）也会使每个参与人获得支付 1。

混合策略均衡中期望支付很低的理由是显而易见的。当他们独立地随机选择时，他们选择不一致从而得到低支付的概率（这里为 4/9）为正并且数值较大。这里对他们最好的就是确定地一起到星巴克。

在这个博弈的纯协调版本中也会有类似的考虑，此时在两个咖啡店里相遇的支付是一样的。此外，两人在均衡中也许会在两个咖啡店间随机选择，但是他们能通过使他们的随机性协调一致而获得更好的结果。假定萨莉与哈里的关系在持续发展中并每天 4 点 30 分在咖啡店相见。他们喜欢变化，因此愿意在两个咖啡店间进行混合。但是，他们都不喜欢事前使选择变得可预测，或者机械性地轮换。（“如果是星期二，就必定在本地咖啡店。”）并且他们不能在 4 点 30 分之前相互联系，这发生在能够瞬间通信的年代以前。（仅仅是几年前，你相信吗?）他们会做什么？一种解决方法为事前双方都同意基于双方均可见的随机信号做出决定。例如，当他们在 4 点 30 分走出各自的房间时，如果在下雨，他们就去星巴克；如果没下雨，就去本地咖啡店。于是，他们就能随机选择并保证能相遇。[1]

借助对约会博弈安全版本混合策略均衡的这些分析，你现在也许能猜出相关的非零和博弈的混合均衡。在前边段落所描绘的纯协调版本中，在两家咖啡店中相遇的支付是相同的，因此混合策略均衡为 $p=1/2$ 与 $q=1/2$。在性别战的变异中，萨莉情愿在本地咖啡店相遇，因为此时她得到支付 2 超过在星巴克相遇时得到的支付 1。她的决定以她关于哈里去星巴克的主观概率是否超过或低于 2/3 为转移。（这里，萨莉的支付与在安全版本中的一样，因此关键的 p 也相同。）哈里则情愿在星巴克相遇，因此他的决定以他关于萨莉去星巴克的主观概率是否超过

或低于 1/3 为转移。因此，混合策略纳什均衡中 $p=2/3$ 和 $q=1/3$。为了提高你的代数技能，我们要求你在本章习题 4 中写出细节。

就像已经看到的，两个坐标轴上的混合概率最优反应分析是一个找到所有纳什均衡的综合的方法。在纯策略的均衡中，有某个策略被采用的概率为 0 的情形发生在方形盒子的边上。真正的混合策略均衡出现在其内部。

如果一个人仅仅希望发现前述博弈的真正的混合策略均衡，那会怎样？这将会比完全最优反应分析更简单。再次考虑表 8—1 与表 8—2 中的博弈的安全版本。找到混合策略均衡的关键在于只有当萨莉关于哈里的选择的主观不确定是正确的时，也就是说，只有当哈里的 p—mix 策略里的 p 值是合适的时候，萨莉才愿意在两种纯策略之间混合。代数学上，这一思想通过解方程 $p=2(1-p)$ 里的均衡 p 值得到证实。这个方程表示当应对哈里的 p—mix 策略时，萨莉从两个纯策略中得到相同的期望支付。在均衡中，方程成立时，就仿佛哈里的混合概率正做着使萨莉无差异的工作。我们强调“仿佛”是因为在博弈中，哈里没有理由使萨莉无差异，其结果仅仅是一个均衡的性质。这个一般性的想法仍然值得记住：在一个混合策略纳什均衡中，每个人的混合概率都使得别人在他的纯策略中无差异。在第 7 章关于零和博弈的讨论中，我们将其称为对手的无差异方法，现在我们发现这一方法在非零和博弈中仍然有效。

8.1.2 懦　夫

最后，再考虑一个非零和博弈。这里考虑第 4 章中参与人试图避免相撞的懦夫博弈。我们将表 4—14 复制为表 8—3。

表 8—3　　懦夫博弈

		迪恩	
		转向（懦夫）	径直向前（勇士）
詹姆斯	转向（懦夫）	0，0	−1，1
	径直向前（勇士）	1，−1	−2，−2

现在，我们引进混合策略。詹姆斯的 p—mix 采用 p 的概率“转向”，$(1-p)$ 的概率“径直向前”，同样地，迪恩的 q—mix 采用 q 的概率“转向”，$(1-q)$ 的概率“径直向前”。为了找到迪恩的最优反应，我们重写了表格如表 8—4，只表示了迪恩的支付并加了一行表示詹姆斯的 p—mix。

表 8—4　　懦夫博弈中的最优反应分析

		迪恩	
		转向	径直向前
詹姆斯	转向	0	1
	径直向前	−1	−2
	p—mix	$p-1$	$3p-2$

迪恩从“转向”中得到的期望支付为 $0\times p+(-1)\times(1-p)=-(1-p)=p-1$。他从“径直向前”中得到的期望支付为 $1\times p+(-2)\times(1-p)=p-2(1-p)=3p-2$。于是，当 $p-1>3p-2$ 或 $2p<1$，或 $p<1/2$ 时，迪恩选择“转向”会更好；如果詹姆斯不倾向“转向”，迪恩选择“径直向前”就会有很大的风险，因此他“转向”会更好。相反，如果 $p>1/2$，迪恩选择“径直向前”会更好；当 $p=1/2$ 时，迪恩从他的两种选择中得到相同的支付，即 $(1/2)-1=3\times(1/2)-2=-1/2$。因此，他就像在$q-$mix的任意组合中无差异一样，在两个选择中也是无差异的。我们将通过画出应对詹姆斯的 $p-$mix 的两个纯策略的两条支付曲线来检验最优反应规则的工作留给读者。

我们将迪恩的最优反应与詹姆斯的最优反应表示在通常的图表中（图 8—3）中。在该图中，我们看到两个博弈的纯策略纳什均衡为最优反应曲线的两个交点。第一点在左上角，此时 $p=0$，$q=1$，因此，詹姆斯选择“径直向前”而迪恩选择“转向”。第二点在右下角，此时 $p=1$，$q=0$，因此，詹姆斯选择“转向”而迪恩选择“径直向前”。我们也发现了第三个交点，即第三个纳什均衡。这一个为混合策略，此时 $p=1/2$，$q=1/2$。在这一点，两人选择的不确定正好能够使得对方无差异。

从图 8—3 中，我们能够计算出迪恩采用他的每个纯策略应对詹姆斯的均衡 $p-$mix 的期望支付。当 $p=1/2$ 时，迪恩从“转向”中得到支付 $p-1=-1/2$；从“径直向前”中得到支付 $3p-2=-1/2$，它们是相等的。这正是詹姆斯的均衡$p-$mix的“对手无差异”性质。于是，迪恩在混合策略均衡中得到支付 $-1/2$。同样，詹姆斯得到支付$-1/2$。如同在 8.1.1 节中考虑的协调博弈的混合策略均衡一样，这些支付很低。这是因为独立地混合能够以正的概率导致一个对彼此都不利的结果——这里就是一次车祸。如果两人能够在重复博弈中达成一个强制性的协议规定双方“转向”或在两个纯策略中交替，他们会得到更好的支付。

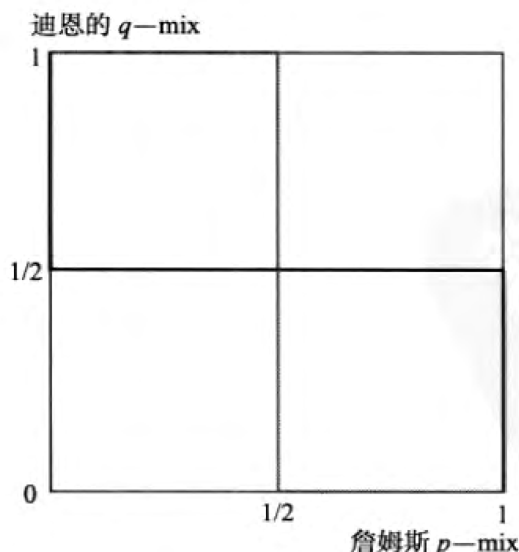


图 8—3　懦夫博弈中的最优反应与纳什均衡

8.2 有三个策略的非零和博弈

为了说明参与人有三个（或更多）策略的非零和博弈的混合策略均衡，我们考虑懦夫博弈的一个扩展版本。在这个扩展版本中，每个参与人不但可以选择是“径直向前”还是“转向”，而且还可以选择转向左边还是转向右边。如果一个参与人转向自己的左边，而另一个参与人转向自己的右边，两人就会相撞，就像他们都选择“径直向前”一样。于是，支付矩阵就如表 8—5 所示：

表 8—5　　　　3×3 懦夫博弈

		迪恩		
		左转	径直向前	右转
詹姆斯	左转	0，0	−1，1	−2，−2
	径直向前	1，−1	−2，−2	1，−1
	右转	−2，−2	−1，1	0，0

首先注意到博弈具有通常的纯策略均衡，此时一个参与人“径直向前”，另一个人“转向”。在这种情况下，朝哪边转向对于他自己而言是无差异的。因此存在四个均衡：（径直向前、左转）与（径直向前、右转），此时迪恩是懦夫；（左转、径直向前）与（右转、径直向前），此时詹姆斯是懦夫。

其次，我们希望找到一个完全的混合策略均衡，在均衡中每个参与人选择所有三种策略的概率都为正。假定詹姆斯的混合概率为选“左转”为 p_1，选“径直向前”为 p_2，选“右转”为 $(1-p_1-p_2)$。同样，写出迪恩的混合概率为 q_1，q_2 与 $(1-q_1-q_2)$。

根据前面章节中的逻辑，我们知道只有当迪恩关于詹姆斯选择的主观不确定恰好是正确的时候，他才会采用混合策略。也就是说，只有当詹姆斯的 p－mix 正好使迪恩在所有的选择中都无差异时，迪恩才会采用混合策略。为了决定 p 的均衡水平，我们必须建立如下等式，即对于詹姆斯的混合策略，迪恩采用三个策略所获得的支付都相等：

$$-p_2-2(1-p_1-p_2)=p_1-2p_2+(1-p_1-p_2)=-2p_1-p_2$$

或

$$-2+2p_1+p_2=1-3p_2=-2p_1-p_2$$

表达式的前两部分得到 $2p_1=3-4p_2$，后两部分得到 $2p_1=2p_2-1$。于是，可以得到 $3-4p_2=2p_2-1$，即 $p_2=4/6=2/3$。运用这一结果，可得 $2p_1=2(2/3)-1=1/3$，所以 $p_1=1/6$。同样地，要使詹姆斯选择混合策略，迪恩的 q－mix 中 q_1、q_2 的均衡值必须使得詹姆斯在他的三个纯策略中无差异，因此 $-2+2q_1+q_2=1-3q_2=-2q_1-q_2$。解这个方程得 $q_2=2/3$，$q_1=1/6$，因此 $(1-q_1-q_2)=1/6$。这里，两个参与人选择“径直向前”的概率都是选择“转向”的概率的两倍。他

们也比在 8.1.2 节中分析的这一博弈的传统版本中更愿意选择“径直向前”，为什么？在懦夫博弈的标准版本中，一个参与人选择“转向”能保证至少得到一个“安全”结果（支付－1）。然而，在懦夫博弈的这个变化版本中，转向左或右都不能保证是安全的，总是有可能存在一次相撞。这个事实使得无论哪种转向都比在标准博弈中缺少吸引力，于是，就使得每个司机在完全混合策略均衡中选择“径直向前”的意愿增加了。

但是，这并不是这个博弈的唯一的混合策略均衡。这里存在另外一个仅仅是局部混合的均衡。也就是说，存在另外的一个混合策略均衡，在这个混合策略均衡中，两个参与人在他们的混合策略中没有完全使用三个策略。

一种可能是参与人在“径直向前”(S) 与其他两个策略中的一个——比如说，“左转”(L) ——混合。为了了解这样一个均衡是如何产生的，假定詹姆斯在他的混合策略中只用 L 与 S，选择 L 的概率为 p，选择 S 的概率为 $(1-p)$。跟从前一样，只有当迪恩采用三个策略应对詹姆斯的混合策略的支付都相等时，迪恩才会在均衡中用混合策略。在这种情况下，迪恩应对詹姆斯的 p－mix 的支付为：

从 L 得到：$0p-1(1-p)$ 或 $p-1$

从 S 得到：$1p-2(1-p)$ 或 $3p-2$

以及

从 R 得到：$-2p-1(1-p)$ 或 $-p-1$

我们可以看到 R 是被 L 与 S 占优的，于是在迪恩的混合策略中不被采用。如果迪恩在混合策略中仅使用 L 与 S，他应该在这两个策略间是无差异的。于是，式 $p-1=3p-2$ 或 $p=1/2$ 一定成立。迪恩在均衡中的期望支付于是为 $-1/2$。并且当迪恩在 L 与 S 间混合时，一个相同的计算表明詹姆斯不会采用他的 R，而且他在 L 与 S 之间的混合概率为 50：50。这个均衡类似 2×2 懦夫博弈中的混合策略结果，这个结果你自己可以计算出来。

然而，另一种可能是只有一个参与人在均衡中使用混合策略。存在这样一个均衡，在均衡中，一个参与人选择纯策略，另一个参与人以正的概率在 L 与 R 之间混合。例如，当詹姆斯选择纯策略 S 时，迪恩在 L 与 R 之间是无差异的，于是愿意在两者之间以任意比率混合，但不想选择自己的 S。反过来，当迪恩在 L 与 R 之间以任意比率混合时，很明显，詹姆斯的最优反应为纯策略 S。于是，我们得到一个纳什均衡，同时，将两个参与人对调就得到另一个同样的均衡。

我们已经知道纯策略越多导致混合策略均衡的可能性越大，包括非唯一的或多个混合策略均衡。我们已经在应用中或是在一些特殊的例子中研究了这些可能性。我们将在下一节中考虑非零和博弈中混合策略博弈的一些一般性理论。

8.3 混合策略均衡的一般性讨论

现在我们已经了解了在零和博弈与非零和博弈中怎样寻找混合策略均衡，值

得思考的是这些均衡的一些其他性质。特别地，在本节中我们将重点介绍混合策略均衡的一些一般性性质，并介绍一些看起来与以前相矛盾的结果，直到你能够完全分析这些有疑问的博弈。

8.3.1 弱均衡

图 7—3 与图 8—2 中的最优反应图形表明，当一个参与人选择自己的均衡混合策略时，另一个参与人的最优反应为包括两个纯策略的极端情况的任意混合。于是，混合策略均衡只是一个弱纳什均衡。当一个参与人选择自己的均衡混合策略时，另一个参与人没有理由背离他自己的均衡混合策略。但是，如果他选择另一个混合策略甚至纯策略，他不一定会得到更坏的结果。只要别的参与人选择正确的（均衡）混合策略，每个参与人在其纯策略以及它们的任意混合中是无差异的。这也是一个混合策略纳什均衡的非常一般性的性质。

这个性质看起来动摇了混合策略纳什均衡作为博弈解的概念的基础。为什么当另一个参与人选择自己的混合策略时，一个参与人应当选择自己合适的混合策略？为什么不简单一些，就选择一个自己的纯策略？毕竟，期望支付是相同的。答案在于如果这样做，就将不是一个纳什均衡，它将不是一个稳定的结果，因为另一个参与人就不会选择自己的混合策略。例如，如果哈里在约会博弈的安全版本中选择纯策略星巴克，于是萨莉可以通过将混合策略 50：50 转为纯策略星巴克，从而在均衡中获得更高的支付（1 代替了 2/3）。

8.3.2 对手的无差异与防止被对手利用

混合策略均衡中的策略选择对于该参与人而言是无差异的，此性质有另外的含义。当一个参与人采用混合策略时，我们知道在另一个人混合策略中出现的纯策略对他自己而言是无差异的，并且我们已经将此称为对手无差异性质。这个性质意味着每个参与人的均衡混合概率都能够通过解对手的无差异方程得到，即令对手的所有纯策略应对混合策略得到的期望支付相等。于是，这形成了第 2 节中讨论的寻找混合策略均衡的一般性方法的基础。

在一个零和博弈中，存在更多的含义。例如，在我们的网球博弈中，当埃弗特得到一个更高的支付时，纳芙拉蒂洛娃得到一个更低的支付。因此，思考混合策略均衡的另一种方式为，当纳芙拉蒂洛娃选择自己的正确的混合策略时，不管埃弗特如何选择，她都会得到相同的结果。任何一个别的选择都将被埃弗特利用，从而成为她的一个优势，成为纳芙拉蒂洛娃的一个劣势。换句话说，每个参与人的均衡混合策略都具有防止被利用（prevent exploitation）的性质。这个性质只属于零和博弈，因为在非零和博弈中，对别的参与人更好并不意味着对自己更坏。但是，在零和博弈中，这个性质给出了一个重要的更为积极的混合的目的。

8.3.3 在零和博弈中混合策略的违反直觉的结果

有混合策略均衡的博弈也许有许多在第一眼看来是违反直觉的性质。它们中最有趣的是博弈支付结构的变化导致的均衡混合策略的变化。为了举例说明，我们再次回到埃弗特与纳芙拉蒂洛娃之间的网球博弈。

假定纳芙拉蒂洛娃致力于提高她防守底线球的技术，使得埃弗特采用 DL 策略应对纳芙拉蒂洛娃防守底线的成功率从 50%降到 30%。纳芙拉蒂洛娃技术的提高改变了表 7—1 中所示的支付表，包括每个参与人的混合策略。在表 8—6 中，我们列出了新表，只表示了行的参与人（埃弗特）的支付。

表 8—6　　　　混合概率中违背直觉的改变

		纳芙拉蒂洛娃		
		DL	CC	q—mix
埃弗特	DL	30	80	$30q+80(1-q)$
	CC	90	20	$90q+20(1-q)$
	p—mix	$30p+90(1-p)$	$80p+20(1-p)$	

与表 7—1 中表格唯一的不同在表格的左上角，这里原先的 50 变为了 30。支付表格中的这个变化并没有为博弈带来一个纯策略均衡，因为参与人仍然有相互冲突的利益，纳芙拉蒂洛娃仍然希望她们的选择能够一致，埃弗特也仍然希望她们的选择不同。我们仍然得到一个将产生混合策略均衡的博弈。

但是，这个新博弈的均衡混合策略与 7.2 节中计算的会有什么不同？一眼看上去，许多人会说纳芙拉蒂洛娃应该多防守 DL，因为她这样做会得到更多支付。于是，假设她均衡的 q—mix 应该赋予 DL 更大的权重，因此其均衡 q 值应该超过以前计算出的 0.6。

但是，当我们用埃弗特的无差异条件计算纳芙拉蒂洛娃的 q—mix 时，我们得到 $30q+80(1-q)=90q+20(1-q)$，或 $q=0.5$。事实上的 q 值为 50%，与许多人直觉预测的大于最初的 q 值 60%相反。

尽管这个直觉看起来是有理由的，但是它遗漏了策略理论的一个重要方面：两个参与人的相互作用。当支付改变之后，埃弗特也会重新评估自己的均衡策略；当纳芙拉蒂洛娃决定自己新的混合策略时，她必须将新的支付结构和埃弗特的行动考虑进去。特别地，由于纳芙拉蒂洛娃现在擅长防守 DL，埃弗特在她的混合策略中就更多地采用 CC。考虑到这一点，纳芙拉蒂洛娃也就会更多地防守 CC。

我们可以通过计算埃弗特新的混合策略来更清晰地了解这一点。她的均衡 p 值必须使得 $30p+90(1-p)$ 与 $80p+20(1-p)$ 相等，于是，她的 p 值为 7/12，等于 0.583，或 58.3%。通过比较新的均衡 p 值与第 7 章中计算得到的 70%，我们发现对于纳芙拉蒂洛娃技术的提高，埃弗特明显地降低了选择 DL 的次数。埃弗

特已经考虑到了她现在面对的对手具有更好的 DL 防守技术的事实，因此，在她的混合策略中，最好少用 DL。凭借这一点，埃弗特使纳芙拉蒂洛娃最好也降低防守 DL 的频率。纳芙拉蒂洛娃的其他混合策略选择，特别是频繁使用 DL 的混合策略，都将被埃弗特利用。

因此，纳芙拉蒂洛娃的技术就白提高了吗？不，但是我们必须正确地评判它——不能通过一个策略使用的频率或别人使用它的频率，而是通过它导致的结果来评判它。当纳芙拉蒂洛娃使用她的新的均衡策略 $q=0.5$，埃弗特采用任何一种纯策略的成功百分比为$(30\times0.5)+(80\times0.5)=(90\times0.5)+(20\times0.5)=55$，它低于原先例子中的成功百分比 62。于是，纳芙拉蒂洛娃的成功百分比也会上升，从 38 升到 45，因此她确实通过提高技术获利。

与考虑纳芙拉蒂洛娃的支付改变时其反直觉的策略反应不同的是，当考虑她的期望支付时，其反应就完全符合直觉了。事实上，当支付改变时，参与人对期望支付的反应不是反直觉的，尽管策略反应是反直觉的。[2] 在参与人的策略反应中的这个反直觉结果的最有趣的方面在于它传递给网球选手（或一般的策略博弈参与人）的信息。这个结果等于说纳芙拉蒂洛娃应该提高她防守底线球的技术以便使得自己更少地防守底线球。

存在可能反直觉结果的别的类似例子，包括在体育运动中经常发生的某些情形。在体育运动中，总是存在一些相对安全的策略，它们即使被对方预料到也不会产生灾难性的失败，但它们即使没有被预料到也不会十分好。另一些策略是冒险的，如果另一方对它无准备，它会非常出色，但是如果另一方有准备，就会悲惨地失败。于是，在橄榄球比赛中，第三次进攻中为了前进一码，中部的直跑是安全的，一个长跑是冒险的。下面的例子就包含“安全对冒险”策略的思想。另外，尽管我们大多数的例子中的支付采用的都是解释性数据，这里我们改变这个习惯来强调问题的普遍性。于是，我们令支付为一般的代数符号，其大小关系仅仅由与需考虑的基本策略相关的情况决定。

考虑任何一个你有两个纯策略的零和博弈。让我们称相对安全的策略为 P，更具风险的策略为 R。对手也有两个我们同样称为 P 与 R 的纯策略。表 8—7 为你博弈成功的概率矩阵，这些不是你的支付。当规定 $a>b>c>d$ 后，“安全”与“冒险”就有它们特定的意义了。如果对方无准备，冒险策略确实非常好（你的成功概率为 a），但是如果对方有准备，就会非常坏（你的成功概率为 d）。安全策略无论在哪种情况下都差不多好（你成功的概率为 b 或 c）。但是，如果对方预料到，会稍微坏一点（$c<b$）。

表 8—7　　冒险与安全策略的成功概率矩阵

		对手的预期	
		P	R
你的选择	P	c	b
	R	a	d

令你成功时的支付或效用为 W，失败时为 L。关键一点在于 W 远远大于 L。注意到 W 与 L 并不一定是钱的数量，因此它们能用来表示第 7 章附录中所解释

的风险厌恶。现在，我们能够写出不同策略组合的期望支付表格，如表 8—8 所示。请注意这个表格是如何构造的。例如，如果你选 P，对手选 R，那么你会以概率 b 得到效用 W，以概率 $(1-b)$ 得到效用 L，你的期望支付为 $bW+(1-b)L$。这个博弈为零和博弈，因此在每个格子中，你对手的支付正好为你的支付的相反数。

表 8—8　　　　冒险与安全策略的支付矩阵

		对手的期望	
		P	R
你的支付	P	$cW+(1-c)L$	$bW+(1-b)L$
	R	$aW+(1-a)L$	$dW+(1-d)L$

在这个混合策略均衡中，你选择 P 的概率 p 由如下式子决定：

$$p[cW+(1-c)L]+(1-p)[aW+(1-a)L]=$$
$$p[bW+(1-b)L]+(1-p)[dW+(1-d)L]$$

这个方程可以简化为 $p=(a-d)/[(a-d)+(b-c)]$。由于 $(b-c)$ 小于 $(a-d)$，所以我们知道 p 接近 1。这正是策略 P 被称为百分比行动的原因，它是在这些情况下的通常行动。而冒险策略 R 只偶尔被采用以使对方保持猜疑，或在橄榄球术语中，称为"使防守老实"。

这个结果有趣的地方在于 p 的表达式独立于 W 和 L。也就是说，理论上你在重要场合下也应该和在小场合中一样以相等的比率混合安全策略与冒险策略。这一点与许多人的直觉相反。他们认为如果场合非常重要时，应该少采用冒险行动。在 10 月的一个普通周日下午，在第三次一码球进攻时发动一个长跑可能会不错，但是如果在超级比赛中这样做就太冒险了。

因此，哪一个是正确的——理论还是直觉？我们猜测读者在这一点上会产生分歧。一些人会认为体育评论员是错误的，他们很高兴发现了一个理论上的论据来驳倒体育评论员的主张。另一些人会站在体育评论员一边，并声称越重大的场合，越要求更安全的行动。还有另外一些人也许会认为当奖金更高时，会采取更冒险的行动，但是他们找不到理论支持，因为理论认为奖金的额度或损失应该对于混合概率是无差异的。

在如前所述的许多理论与直觉产生差异的情况下，我们认为差异要么仅是表面的，要么是理论不够一般化，或者说未丰富到可以涵盖产生直觉的情形的所有特性，才出现分歧，并且应完善理论来消除分歧。这里的问题不一样，它是混合策略中计算支付的基础，就如同概率权重平均或期望支付一样。并且，几乎所有的现存博弈理论都以此为起点。[3]

8.3.4　非零和博弈中另一个违反直觉的结果

这里，我们考虑一个普通的 2×2 非零和博弈，支付表如表 8—9 所示。在实

际博弈中，支付应该是实际的数字，策略也应有特定的名字。在这个例子中，我们再次用一般的代数符号来表示支付，这样我们能够检验出均衡混合策略的概率是如何依赖于这些支付数字的。同样地，我们用一般的抽象文字来表示策略。

表 8—9　　2×2 非零和博弈的一般代数支付矩阵

		列	
		左	右
行	上	a，A	b，B
	下	c，C	d，D

假定博弈有一个混合策略均衡，其中行参与人以概率 p 选择“上”，以概率 $(1-p)$ 选择“下”。为保证列参与人在均衡中也使用混合策略，行参与人的 $p-$mix 必须使得列参与人在两个纯策略“左”与“右”之间是无差异的。为使得列参与人采用两个纯策略应对行参与人的混合策略的期望支付相等，我们得到[4]

$$pA+(1-p)C=pB+(1-p)D，\text{或}$$
$$p=(D-C)/[(A-B)+(D-C)]$$

关于 p 的表达式，令人吃惊的不在于它包含了什么，而在于它未包含什么。注意行参与人自己的支付为 a、b、c 或 d，行参与人的均衡混合概率完全独立于他自己的支付！

同样地，列参与人选择“左”的均衡混合概率 q 也通过下式得到

$$q=(d-b)/[(a-c)+(d-b)]$$

列参与人的均衡混合概率同样完全独立于他自己的支付。

如果你还记得对手无差异的一般性准则，这些结果中令人惊异与违反直觉的部分就能够得到解释了。由于每个参与人的混合概率都通过使对手在他的纯策略间无差异的方式解得，很自然，这些概率就应该取决于对手的支付，而不是自己的支付。但是，同样也要记住只是在零和博弈中，参与人才有真正的理由使对手保持无差异。对手对他的任意一个纯策略的偏爱都将会成为自己的劣势，在非零和博弈中，对手的无差异不具备这样很强的目的性，它仅仅是混合策略均衡的一个逻辑性质。

8.4　非零和博弈中混合策略的实证

在第 7 章中，我们注意到关于混合策略博弈的实验室实验经常发现结果与理论上的均衡不一致，这一点对零和博弈特别正确。正如我们已经看到的，在这类博弈中，每个参与人的均衡混合策略使对手在纯策略中无差异的性质是均衡的一个逻辑性质。不像零和博弈，通常非零和博弈中的参与人没有明确的理由使别的

参与人保持无差异。因此，混合策略计算背后所隐藏的推理对于参与人而言，就更加复杂和难以理解，这在他们的行动中表现了出来。

在一群非零和博弈的实验对象中，我们也许会看到一些参与人采取一个纯策略，而另一些则选择另外一个。这种人群中的混合，尽管不符合混合策略均衡的理论，却有着一个有趣的进化解释，我们将在第 13 章中对这一点进行检验。

另一个实验问题涉及进行多次博弈的对象。收集关于非零和博弈的证据时，循环或随机化每个对象的对手以避免在重复博弈中形成潜在的合作是重要的。在这种实验中，参与人在逐一改变自己的行动。但是，即使我们认定它为真实混合策略的证据，当参与人的支付改变时，他们的混合概率也会改变。根据 8.3.4 节中的理论，这不应该发生。

综上所述，可得出的结论为非零和博弈中的混合策略均衡最好是经过谨慎考虑之后再解释或应用。

8.5 三个或更多策略中的混合：一般性理论

我们通过第 7 章、第 8 章中各种例子传达的思想来为本章总结出一些关于混合策略均衡的一般性理论。这种一般性理论不可避免地需要运用一些代数知识与抽象的思维。对数学无准备或感到厌恶的读者可以跳过这一节，而不会影响阅读的连贯性。

假定行参与人有可用的纯策略 R_1，R_2，…，R_m，列参与人有纯策略 C_1，C_2，…，C_n。将行参与人从策略组合（i，j）中获得的支付记为 A_{ij}，列参与人从策略组合（i，j）中获得的支付记为 B_{ij}，这里 i 取值从 1 到 m，j 取值从 1 到 n。我们允许每个参与人混合可用的纯策略。假定行参与人的 $p-$mix 中概率为 P_i，列参与人的 $q-$mix 中概率为 Q_j。所有的概率必须非负，并且每个集合必须加起来等于 1。于是，有

$$P_1+P_2+\cdots+P_m=1=Q_1+Q_2+\cdots+Q_n$$

我们令 V_i 为行参与人采用纯策略 i 应对列参与人的 $q-$mix 的期望支付。根据我们在几个例子中已经看到的理由，得到

$$V_i=A_{i1}Q_1+A_{i2}Q_2+\cdots+A_{in}Q_n=\sum_{j=1}^{n}A_{ij}Q_j$$

这里，右边的最后一个表达式采用了数学的求和符号。当行参与人采用他的 $p-$mix，而正对应了列参与人的 $q-$mix 时，行参与人的期望支付为

$$P_1V_1+\cdots+P_mV_m=\sum_{i=1}^{m}P_iV_i=\sum_{i=1}^{m}\sum_{j=1}^{n}P_iA_{ij}Q_j$$

行参与人选择他的 $p-$mix 以最大化这个表达式。

同样地，令 W_j 为列参与人采用纯策略 j 应对行参与人的 $p-$mix 的期望支

付，得到

$$W_j = P_1B_{1j} + P_2B_{2j} + \cdots + P_mB_{mj} = \sum_{i=1}^{m} P_iB_{ij}$$

用混合策略应对混合策略，我们得到列参与人的期望支付为

$$Q_1W_1 + \cdots + Q_{nW_n} = \sum_{j=1}^{n} Q_{jW_j} = \sum_{i=1}^{m}\sum_{j=1}^{n} P_iB_{ij}Q_j$$

列参与人选择他的 q—mix 以最大化这个表达式。

每个参与人都同时在假定别人选择最优混合策略时选择他的最优混合策略，这样我们就得到一个纳什均衡。也就是说，行参与人的 p—mix 应该是对列参与人的 q—mix 的最优反应，反过来也成立。让我们从发现行参与人的最优反应规则开始，也就是说，让我们临时固定列参与人的 q—mix，考虑行参与人对 p—mix 的选择。

假定对于列参与人给定的 q—mix 策略，行参与人有 $V_1>V_2$。于是，行参与人能够通过将 R_2 的一些概率转到 R_1 来提高期望支付。也就是说，行参与人减少他选择 R_2 的概率 P_2，增加选择 R_1 的同样数量的概率 P_1。因为 V_1 与 V_2 的表达式不包括任何概率 P_i，因此不管 P_1、P_2 的原始数值是多少，都是正确的。于是，行参与人应该尽可能降低选择 R_2 的概率 P_2，即使它等于零。

这个思想可以立即推广。行参与人应该以降序排列 V_i。在其顶端也许正好存在一个策略，这种情况下应该只有唯一的策略可用，即行参与人应该用一个纯策略。也有可能在其顶端存在两个或更多的策略，这种情况下行参与人应该仅仅在这些策略间混合而不采用其他策略。[5] 当存在这样一个约束时，这些策略间的所有混合使得行参与人获得相同的期望支付。于是，单独的考虑不能够确定行参与人的均衡 p—mix。我们随后会说明列参与人的无差异条件如何完成这项工作，虽然第一眼看起来比较奇怪。

同样的讨论也可应用于列参与人。他应该只采用给了他最高 W_j 的纯策略，或者只在那些使 W_j 位于顶端的纯策略 C_j 间混合。如果存在这样一个约束，那么所有的混合对于列参与人都是同样好的；混合策略的概率不能被单方面考虑所确定。

通常，对于我们在列参与人的 q—mix 中固定的（Q_1，Q_2，…，Q_n）的大多数取值，行参与人的 V_1，V_2，…，V_m 在顶端没有任何约束，于是行参与人的最优反应应该是一个纯策略。反过来，对于行参与人的 p—mix 中固定的（P_1，P_2，…，P_m）的大多数取值，列参与人的最优反应也是一个纯策略。例如，我们在第 7 章、第 8 章中的例子中几次看到的，在图 7—3、图 8—2 与图 8—3 中，对于行参与人的 p—mix 中 p 的大多数取值，列参与人的最优的 q 不是 0 就是 1，反之亦然。行参与人 p 的唯一关键值是应对列参与人的混合策略（选择 0 到 1 之间的任意 q）的最优 p 值，反之亦然。

所有这些条件（在顶端约束，并且别的策略只能得到更坏的结果）构成了一个等式与不等式的复杂集合，这些等式与不等式同时被满足时，就定义了博弈的混合策略纳什均衡。为了更好地理解这一点，假定目前我们已经做了所有工作并

找到了在均衡混合策略中所用到的那些策略。我们总是能够重新标记这些策略使得当列参与人用前 h 个纯策略 C_1，C_2，…，C_h，并且不采用余下的（$n-h$）个纯策略 C_h+1，C_h+2，…C_n 时，行参与人用前 g 个纯策略，即 R_1，R_2，…，R_g，并且不采用余下的（$m-g$）个纯策略 R_g+1，R_g+2，…，R_m。记 V 为行参与人的顶端期望支付 V_i 的约束值，同时，记 W 为列参与人的顶端期望支付 W_i 的约束值。于是，这些等式与不等式可以表达如下。首先，对于每个参与人，我们设那些没被采用的策略的概率为 0，那些被采用的策略的概率之和等于 1：

$$P_1+P_2+\cdots+P_g=1, P_{g+1}=P_{g+2}=\cdots P_m=0 \tag{8.1}$$

并且

$$Q_1+Q_2+\cdots+Q_h=1, Q_{h+1}=Q_{h+2}=\cdots Q_n=0 \tag{8.2}$$

接下来，我们设行参与人对于他所采用的纯策略的期望支付等于顶端约束值：

$$V=A_{i1}Q_1+A_{i2}Q_2+\cdots+A_{ih}Q_h \quad i=1,2,\cdots,g \tag{8.3}$$

并且注意到来自他的未采用策略的期望支付必定更小（这是它们未被采用的原因）：

$$V>A_{i1}Q_1+A_{i2}Q_2+\cdots+A_{ih}Q_h \quad i=g+1,g+2,\cdots,n \tag{8.4}$$

接下来，我们对列参与人做同样的处理，记 W 为他的顶端约束支付值：

$$W=P_1B_{1j}+P_2B_{2j}+\cdots+P_gB_{gj} \quad j=1,2,\cdots,h \tag{8.5}$$

以及

$$W>P_1B_{1j}+P_2B_{2j}+\cdots+P_gB_{gj} \quad j=h+1,h+2,\cdots,n \tag{8.6}$$

为了找到均衡，我们把对 g 和 h 的选择与概率 P_1，P_2，…，P_g；Q_1，Q_2，…，Q_h 当做未知数进行整体求解。

穷举法总很适用。尝试 g 和 h 的特殊选择。也就是，选择一个纯策略的特殊集合作为均衡中采用的候选者。于是，将方程（8.1）与方程（8.5）作为一组 $h+1$ 个同时发生的线性方程，其中 $P_1,P_2,\cdots,P_g$ 与 W 为（$g+1$）个未知数，对它们求解，并检查所有的解是否满足不等式（8.6）。同样地，将方程（8.2）与方程（8.3）作为一组 $g+1$ 个同时发生的线性方程，其中 $Q_1,Q_2,\cdots,Q_h$ 与 V 为（$h+1$）个未知数，对它们求解，并检查所有的解是否满足不等式（8.4）。如果所有条件都满足，我们就已经找到了均衡。如果不满足，再选一组纯策略作为候选者，再试一次。只存在有限次这种选择：行参与人在他的混合策略中可用的纯策略存在（2^m-1）种可能选择；列参与人在他的混合策略中可用的纯策略存在（2^n-1）种可能选择。于是，这个过程经过有限次尝试之后一定会成功结束。

当 m 与 n 比较小时，穷举法是可解决问题的。尽管如此，在每一个特殊问题的计算过程中也会存在捷径。于是，在我们的罚点球博弈的第二个变化中，求解所有不能使用的策略的尝试会告诉我们哪些策略在下一次尝试中应该舍弃。

然而，对于更大一点的问题，基于穷举法或一些特殊方法的解会变得非常复杂。这时就需要更加系统的计算机搜寻算法。这些计算机算法所做的就是同时对两个线性最大化问题（或者是决策理论术语中的线性规划问题）搜寻解：给定一

个 $q-\text{mix}$ 与所有的 V_i 值，选择一个 $p-\text{mix}$ 最大化行参与人的期望支付 $\sum_i P_iV_i$ ，以及给定一个 $p-\text{mix}$ 与所有的 W_i 值，选择一个 $q-\text{mix}$ 最大化列参与人的期望支付 $\sum_j Q_jW_j$ 。然而，对于一个特定的 $q-\text{mix}$，所有的 V_i 值都将是不相等的。如果列参与人在实际博弈中采用这个 $q-\text{mix}$，行参与人将不会采用混合策略而只采用给他最高 V_i 值的那个纯策略。但是，在我们的数值计算方法中，我们不应该用这种极端的方式调整行参与人的策略。如果我们这样做，那么在我们算法的下一步，列参与人的最优的 $q-\text{mix}$ 同样也会极端地改变，并且行参与人所选择的纯策略将不再像看起来那么好。与此相反，算法应该采用一个更加渐进的步骤，一点一点地调整 $p-\text{mix}$ 来提高行参与人的期望支付。于是，对于行参与人新采用的 $p-\text{mix}$，就应该一点一点调整列参与人的 $q-\text{mix}$ 来提高他的期望支付。于是，再回到在 $p-\text{mix}$ 中的另一个调整。这个过程一直持续到再也不能够找到新的提高，那就是均衡。

我们并不需要这个程序的细节，但是我们上面提出的一般的思想已经告诉了我们许多关于均衡的东西。这里有一些重要的经验。

（1）我们从方程（8.1）与方程（8.5）中求解行参与人的均衡混合概率 P_1，P_2，…，P_g。方程（8.1）仅仅增加了概率约束的条件。实质的方程为公式（8.5），它给出了一个条件，这一条件要求列参与人从他所采用的应对 $p-\text{mix}$ 的所有纯策略中都可以得到相同的支付。当列参与人只在意他自己的支付而不在意行参与人的支付时，这看上去似乎有些令人费解，即列参与人要调整他自己的组合，目的是为了与行参与人的选择无异。而事实上，这种困惑也只是表面上的。我们通过考虑行参与人在只关心他自己的支付的动机下对 $q-\text{mix}$ 做出的选择就可以得出方程（8.5）的条件。我们主张列参与人应该只用能够使他针对行参与人的 $p-\text{mix}$ 得到最好支付的那些策略。尽管表面上看起来仿佛行参与人故意地选择 $p-\text{mix}$ 使列参与人无差异，而产生这种结果的真实力量是列参与人自己的有目的的策略选择。

在第 7 章中，可利用解每个参与人的无差异条件构成的方程组从而求出其他参与人的均衡混合策略，这一思想被命名为“对手无差异准则”。我们现在已分别在普通博弈、零和博弈与非零和博弈中对该准则进行了验证。

（2）然而，在零和博弈中，每个参与人都选择他的混合策略使别人无差异的思想并不只是理论上的。存在许多真实的理由使得一个参与人应该按照这种方式行动。若博弈是零和的，在两个参与人的支付之间有一种自然的联系：对于任何 i 与 j，$B_{ij}=-A_{ij}$，并且在所有组合与期望支付之间也存在类似的关系。于是，我们能够将方程（8.5）、方程（8.6）乘以 -1 作为行参与人的支付而不是列参与人的支付。我们将这个条件的这些“零和版本”写作方程（8.5z）与方程（8.6z）：

$$V=P_1A_{1j}+P_2A_{2j}+\cdots+P_gA_{gj}\quad j=1,2,\cdots,h \tag{8.5z}$$

与

$$V<P_1A_{1j}+P_2A_{2j}+\cdots+P_gA_{gj}\quad j=h+1,h+2,\cdots,n \tag{8.6z}$$

注意，将方程（8.6）乘上－1 变为方程（8.6z）。方程（8.6z）颠倒了不等式的方向。

方程（8.5）与方程（8.6）告诉我们只要行参与人采用他的均衡混合策略，列参与人（在这个零和博弈中，行参与人也是一样）在均衡中采用的任一个纯策略其支付都是相等的。列参与人通过选择那些策略之一不能为自己谋得更多的利益，即，在零和博弈中不能对行参与人造成任何损害。不仅如此，方程（8.6z）告诉我们列参与人采用任何别的策略，行参与人都会更好。换句话说，这些条件告诉我们行参与人的均衡混合策略不能被列参与人利用。于是，我们在更通常的背景下看到零和博弈中混合策略的目的性极强的角色，正如我们在第 7 章中看到的。我们也更清楚地看到为什么它只在零和博弈中起作用。

（3）现在，我们回到更普通的非零和博弈情形。注意到系统包含的方程（8.1）与方程（8.5）有（$h+1$）个线性方程与（$g+1$）个未知数。通常，如果 $h>g$，这样一个系统无解；如果 $h=g$，正好有一个解；如果 $h<g$，有许多解。反过来，包含方程（8.2）与方程（8.3）的系统有（$g+1$）个线性方程与（$h+1$）个未知数。通常，如果 $g>h$，这样一个系统无解；如果 $g=h$，正好有一个解；如果 $g<h$，有许多解。因为在均衡中，我们希望两个系统都满足，通常我们需要 $g=h$。于是，在一个混合策略均衡中，两个参与人采用同样数目的纯策略。

我们一直说“通常”是因为系数的偶然组合与方程的等式右边会导致例外。特别地，方程极多、未知数极少以致没有解是可能的。这正是在 7.4.2 节中提到的“例外情形”中所发生的。

（4）我们观察到在所采用的策略与它们的支付之间有一种非常特殊的关系。行参与人以正的概率采用 P_1 到 P_g 的策略，并且方程（8.3）表明这些纯策略中的任何一个被用来应对列参与人的均衡混合策略，他正好得到支付 V。对于行参与人剩余的纯策略，方程（8.4）表明当它们被用来应对列参与人的均衡混合策略时，得到的支付低于 V，于是，它们也就不被采用，即 P_h+1 到 P_m 都为 0。换句话说，对于任意 i，既要 $V>A_{i1}Q_1+A_{i2}Q_2+\cdots A_{ih}Q_h$，又要 $P_i>0$，是不可能的。

至少这些不等式中的一个必须变为等式。这就是被所谓的互补松弛（complementary slackness）的准则，正如在数学优化（规划）中一样，它在博弈与均衡的普通理论中也非常重要。

（5）回到零和博弈情形。当两个参与人都选择他们的均衡混合策略时，行参与人的期望支付为

$$V=\sum_{i=1}^{m}\sum_{j=1}^{n}P_iA_{ij}Q_j,$$

并且列参与人的支付正好是这个支付的负数。此外，当行参与人对于给定的 $q-$mix，选择他的 $p-$mix 以最大化这个表达式，而且同时，列参与人对于给定的$p-$mix，选择他 $q-$mix 的以最大化这个表达式的负数或最小化这个表达式时，均衡就出现了。如果我们将这个表达式作为 P_i 与 Q_i 的函数，我们将它画在一个足够高的维度的空间中，它看起来像一个马鞍。前后横切这个马鞍的切片看起来

像一座山谷或一个 U 字，它的最小点在中部。而从侧面横切马鞍的切片看起来像一个山顶或一个倒 U，它的最大点也在中部。如果每个参与人都正好有两个纯策略，$p-\text{mix}$ 与 $q-\text{mix}$ 就能被一个单独的数字描述，即选择第一个策略的概率。（对于每个参与人，选择第二个纯策略的概率为 1 减去选择第一个纯策略的概率。）我们于是能够在三维空间中画出这个图形，这里 x 轴与 y 轴在水平面上，z 轴垂直向上。$p-\text{mix}$ 用 x 轴表示，$q-\text{mix}$ 用 y 轴表示，数值 V 用 z 轴表示。这个马鞍形表面沿着 x 轴的横切面将表示出关于 p 的 V 的最大值，于是是一个山顶。并且沿着 y 轴的横切面将表示出关于 x 的 V 的最小值，于是是一座山谷。于是，这个图形看起来像一个马鞍，如图 8—4 所示。这样一个均衡被称为一个**鞍点**（saddle point）。

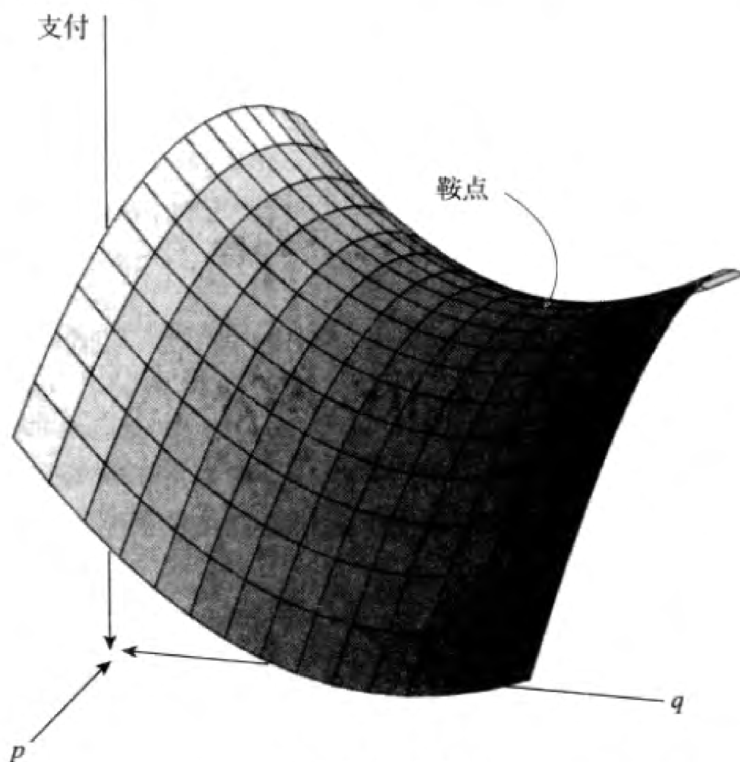


图 8—4　鞍点

均衡中的 V 值——也就是关于 p 的最大值，关于 q 的最小值——被称为零和博弈的最小最大值。这样一种均衡的思想，与定义它的方程（8.5z）、方程（8.6z）那样的条件方程一道，是第一个重要的博弈理论成果，该成果出现在冯·诺依曼与摩根斯坦 20 世纪 40 年代的论文之中。这被称为他们的最小最大定理。

8.6 总　结

在非零和博弈中，当参与人关于别人行动的主观不确定都正好在一个合适的水平上时，混合策略就会产生。在均衡中，一个参与人选择他的均衡混合策略就

仿佛是在让对手在纯策略之间保持无差异。最优反应分析提供了决定均衡混合概率的框架。在大型博弈中，参与人有三个（或更多）策略，混合策略均衡也许是完全或部分混合，并且有可能存在多个或非唯一均衡。

混合策略均衡仅仅是一种弱均衡，因为当参与人的对手采用混合策略时，参与人在所有可能混合策略中是无差异的。在零和博弈中，对手无差异性质能够被解释为帮助参与人防止被对手利用。并且违反直觉的结果——参与人对那些支付上升的策略采用的概率较通常认为的要小，并且在非零和博弈中，均衡混合策略仅仅取决于其他参与人的支付——在零和博弈与非零和博弈中都会出现。实验证据表明理论的预言应该被谨慎使用。但是，本章与第 7 章中的理论能够被推广到任何大小与类型的博弈中。

关键术语

互补松弛（complementary slackness）

防止被利用（prevent exploitation）

鞍点（saddle point）

主观不确定（subjective uncertainty）

习　题

1. 采用最优反应分析方法在下面博弈的混合策略中寻找一个混合均衡。画出最优反应图，并在图上标明均衡混合策略，同时说明均衡中每个参与人的期望支付。

		列	
		左	右
行	上	4，0	−1，2
	下	1，1	2，−1

2. (a) 找出下面非零和博弈的所有纯策略纳什均衡。
   (b) 找出这个博弈中的混合策略均衡。在这个均衡中，参与人的期望支付是什么？

		列			
		A	B	C	D
行	1	1，1	2，2	3，4	9，3
	2	2，5	3，3	1，2	7，1

3. (a) 下表为一个两人同时行动博弈的金钱支付。找出这个博弈的混合策略纳什均衡以及这个均衡中参与人的期望支付。

(b) 当行参与人选择“下”时，两个参与人共同获得最多的金钱。然而，在均衡中，行并不总是选择“下”，为什么？你能找到一个方法能够产生一个更合作的结果吗？

		列	
		左	右
行	上	1，16	4，6
	下	2，20	3，40

4. 表 4—13 给出了哈里与萨莉之间约会博弈性别战版本。
   (a) 采用最优反应分析求解博弈中的混合策略均衡。
   (b) 确定均衡中每个参与人的期望支付并将这些支付与博弈的两个纯策略纳什均衡的支付相比较。参与人怎样才能比他们随机独立行动做得更好？
5. 回忆第 4 章习题 7，一个老妇人寻求帮助横过马路。只需要一个人就可以帮助她，更多的人也可以但结果并不比一个人好。你和我是在附近能够提供帮助的两个人，我们分别同时选择是否这样做。我们中的每个人都会从老妇人成功过马路中获得愉悦，价值为 3（不管谁帮助了她）。但是帮助她的人将会付出成本 1，这是我们提供帮助时花费的时间的价值。要求你将此构造成一个博弈并写出支付表。如果你做了这道题，你可能已找到这个博弈所有的纯策略纳什均衡。现在，请找出这个博弈的混合策略均衡。
6. 考虑下面这个懦夫博弈的变形版本，当迪恩为“懦夫”时，詹姆斯成为“勇士”的支付为 2，而不是 1。

		迪恩	
		转向（懦夫）	径直向前（勇士）
詹姆斯	转向（懦夫）	0，0	−1，1
	径直向前（勇士）	2，−1	−2，−2

   (a) 找出这个博弈的混合策略均衡，以及参与人的期望支付。
   (b) 将这个结果与 8.1.2 节中原始博弈的结果进行比较。迪恩选择“径直向前”的概率比以前更高了吗？詹姆斯选择“径直向前”的概率是多少？两个参与人的期望支付是多少？按照新的支付结构，这些均衡结果之间的差异荒谬吗？按照对手无差异原则对你的结论进行解释。
7. 回忆第 4 章习题 11 中的博弈。三个竞争者（A、B 与 C）参与一个博弈，其奖品价值 30 美元。每个人可以买一张价值 15 美元或 30 美元的票，也可以不买票。他们同时独立地做出选择。于是，根据票的购买决策，博弈的组织者颁发奖品。如果没有人买票，奖金就不颁发。如果只有一个人买了一张票，这个人就得到奖品。若有两个或更多的人买了票，如果只有一个人买的票成本最高，则这个人得到奖品；如果两个或三个竞争者买了最高成本的票，奖品就在他们之间平分。找出这个博弈的混合策略对称的均衡，此时，每个竞争

者都在他们的购票选择中混合，并且三个竞争者的混合概率也是相同的。

8. 回忆第 6 章习题 5 中海滩上的冰淇淋小贩的博弈。构造一个博弈的 5×5 表格，并找出混合策略纳什均衡。解释在均衡混合策略中，为什么一些纯策略未被采用并证明（通过对手的均衡混合策略）它们就不应该被采用。

9.（选做）回忆第 4 章习题 12 中电影《美丽心灵》中的酒吧场景。这里，我们考虑当有 $n>2$ 个年轻人参与时该博弈的混合策略均衡。

   (a) 从考虑所有 n 个年轻人都以概率 P 追求唯一的金发碧眼女生的对称情形开始，给定其他人都采用混合策略，这个概率就被每个年轻人在金发碧眼女生与浅黑肤色女生之间无差异这一条件所决定。什么条件保证了每个参与人无差异？这个博弈中，均衡的 P 值是多少？

   (b) 这个博弈也存在不对称的混合策略均衡。在这些均衡中，有 $m(m<n)$ 个年轻人以概率 Q 追求金发碧眼女生，余下的 $n-m$ 个年轻人追求浅黑肤色女生。给定其他人的选择，什么条件保证了 m 个年轻人无差异？什么条件必须成立才能使余下的 $n-m$ 个参与人不想偏离选择追求一个浅黑肤色女生纯策略？在这个不对称均衡中，均衡的 Q 值是多少？

【注释】

[1] 为了严格地分析这种情形，我们应当将关系在持续发展中、渴求变化等因素明确地引入博弈的结构及其支付中。我们希望你能够接受这个简单的故事，因为它能很好地解释所有参与人都可察觉的随机机制的作用。

[2] 关于改变一个特定单元格中的支付对均衡混合策略与均衡中期望支付的影响的一般性理论，参见 Vincent Crawford and Dennis Smallwood, "Comparative Statics of Mixed-Strategy Equilibria in Noncooperative Games," *Theory and Decision*, vol. 16, May 1984, pp. 225 -232。

[3] 参见 Vincent P. Crawford, "Equilibrium Without Independence," *Journal of Economic Theory*, vol. 50, no. 1, February 1990, pp. 127 - 154; James Dow and Sergio Werlang, "Nash Equilibrium Under Knightian Uncertainty," *Journal of Economic Theory*, vol. 64, no. 2 (December 1994), pp. 305 - 324。上述两篇论文是启迪另一种博弈理论基础的论文。而我们在本书第一版本中对此问题的揭示启发了一篇将新方法运用于此的论文：Simon Grant, Atsushi Kaji, and Ben Polka, "Third Down and a Yard to Go: Recursive Expected Utility and the Dixit-Skeath Conundrum," *Economic Letters*, vol. 73, no. 3, December 2001, pp. 275 - 286。非常不幸，它所用到的先进的概念已经超出了本书介绍水准所允许的程度。

[4] 由于存在一个混合策略均衡，概率 p 必须在 0 到 1 之间。这就需要 $(A-B)$ 与 $(D-C)$ 有相同的符号。如果 A 大于 B，D 就必须大于 C；如果 A 小于 B，D 就必须小于 C。（否则，纯策略左和右之一会占优另一个。）

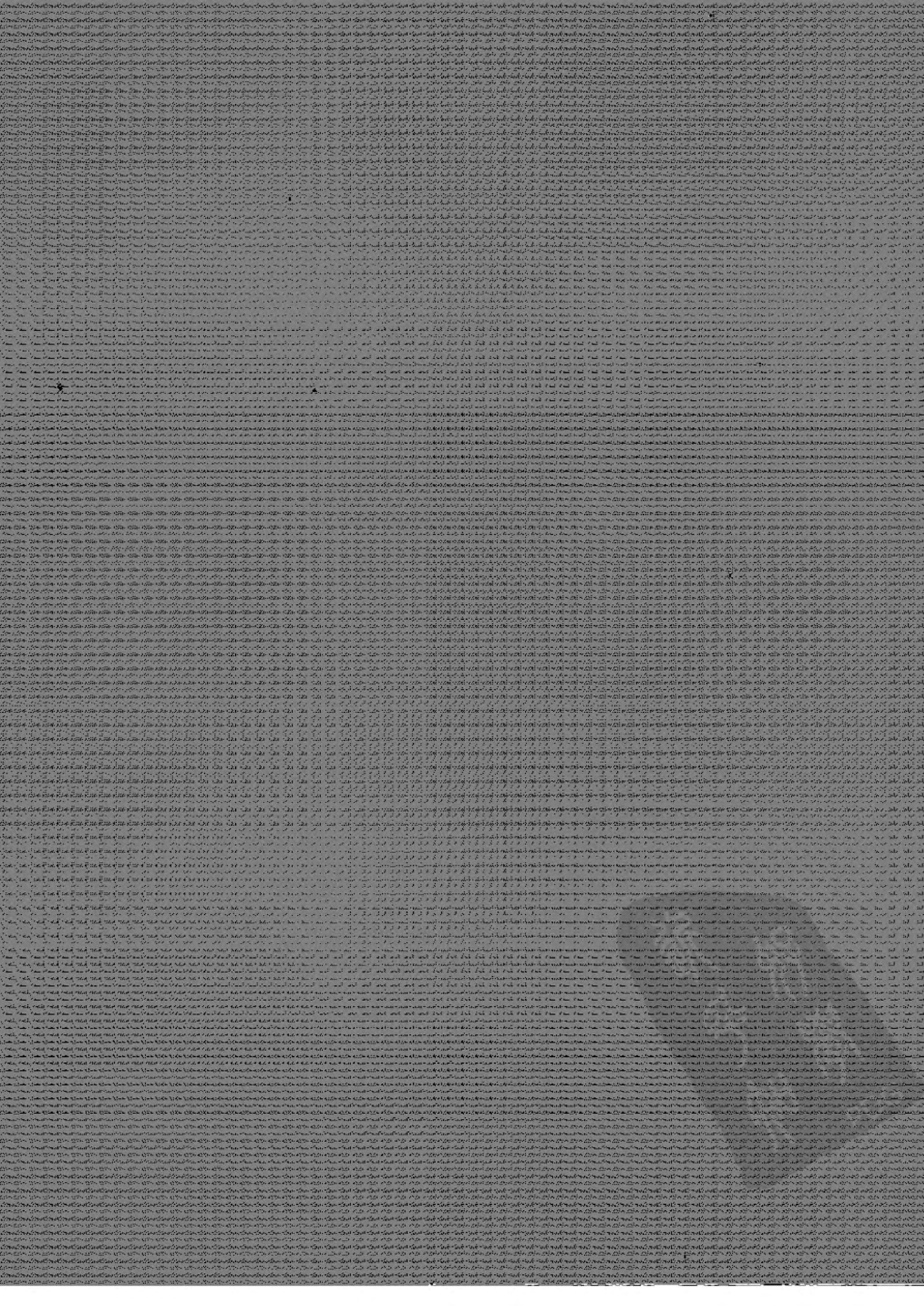
[5] 在数学术语中，表达式 $\sum_i P_i$ 关于 P_i 是线性的。于是，它的最大值必定位于 P_i 的可允许集合的一个极值点上。

经济科学译库·策略博弈 经济科学译库·策略博弈 经济科学译库·策略博弈 经济科学译库·策略博弈

第三部分

某些更为广拓的博弈与策略类型





第 9 章　不确定性与信息

我们曾研究过的许多博弈中，都存在着结果的不确定性。有时，不确定性是参与人混合策略的结果，但更多的时候则是博弈本身的一些特性。比如，网球或足球比赛中，我们假设参赛者仅知道不同策略组合成功的概率，但不能精确预测这些策略运用的结果。在某些博弈中，每个人都无法确切知晓其他人的动机和支付；这里的不确定性是非对称的，因为每个人都非常清楚自己的支付。在本章中，我们将发展出不同的思想和方法来分析这种存在非对称不确定性的博弈。

特别地，我们将注意力转向博弈的不同参与人拥有不同数量信息的情形。正如第 2 章指出的，操纵信息是一个重要的策略维度。我们将发现信息何时能何时不能通过可置信的方式进行交流。我们也将考察传达或隐匿己方信息以诱导他人信息的谋略。此类谋略在第 1 章、第 2 章中被简称为甄别（screening）和信号传递（signaling）。本章将更深入细致地讨论它们。

对博弈中信息及其操纵的研究，在最近 25 年非常活跃且深受重视。经济学中诸多令人困惑的问题都因对信息的研究而变得明朗，比如激励合约、企业组织、劳动力市场、耐用品市场、政府对商业的规制等。[1] 近年来，同样的概念已经被政治科学家用于解释财政收支政策变化与选举的关系、议会选举等诸如此类的问题。这一思想也延伸到了生物学，进化博弈认为孔雀漂亮的大尾巴是一个信号。本章将向你介绍这些思想，以便你日后深入了解一些令人着迷的最新发展。也许更重要的是，你将认识到信号传递和甄别在日常生活博弈中的重要性，并因此提高博弈技巧。

9.1 非对称信息：基本思想

许多博弈中，对发生了什么或者将要发生什么，一些参与人比另一些参与人确实知道得更多。这种信息优势，或者非对称信息，在现实的策略博弈情形中屡见不鲜。在最基本的层面上，每个参与人都非常了解自己的偏好和支付——比如边缘政策博弈中对风险的态度、讨价还价中的耐心、国际关系中和平或好战的意图，但对其他参与人的偏好和支付却知之甚少。同样，参与人对自己的禀性特征（比如员工的技能、汽车或健康保险申请人的冒险行为等）是了解的，而对其他人的禀性特征却不太了解。有时，某个参与人可采取的某些行动，比如一个国家的武力和战备，并不为另一方完全知悉。即便知道了各种可能的行动，另一方也可能观察不到一方事实上采取的行动，比如，经理并不能准确地获悉员工努力和勤奋的程度。最后，某些事实上的结果（比如一个家庭在洪水或地震中遭受的损失的真实货币价值）可以被一方观察到但却不能被另一方观察到。

通过操纵他人对己方能力和偏好之了解，就可以影响博弈的结果。因此这种对非对称信息的操纵本身就是一个策略博弈。人们可能认为，每个人都总是想隐匿自己的信息并套出他人的信息，事实并非如此。下面以例子列出了各种可能性。信息优势方可能会采取如下方式之一：

（1）隐匿信息或披露误导信息。零和博弈中采取混合行动，是因为你不想让他人知悉你的行动；扑克赛中虚张声势，是因为你想误导对手的猜牌。

（2）诚实地披露精心挑选过的信息。在策略行动时，你希望对手看到你已经采取的行动，以便他们做出你所希望的反应。例如，如果在一种紧张的情形里你并没有恶意，则你会希望其他人能相信你的善意，以免引起不必要的纷争。

同样，信息劣势方可能采取如下方式之一：

（1）套出信息或从谎言中筛选出真实信息。比如，老板想知道新员工的能力以及目前员工的努力程度，保险公司想知道客户的风险程度、索赔人的损失程度以及任何有助于减少其责任的索赔人的疏忽。

（2）保持无知。当你不知道对手的策略行动时，你就可以避免他的承诺或威胁。例如，高层的政客或管理者常常受益于“可置信的拒绝力”（credible deniability）。

传达信息最简单的方式似乎是直接告诉他们；同样，套出信息的最简单方式似乎是直接询问。然而，策略博弈中，参与人应当意识到其他人可能不会讲真话，而他们自己的主张也可能不会被其他人相信。也就是说，口头几句话的**置信度**（credibility）是值得怀疑的。我们会在本章下一节讨论直接沟通在何时才可信。

我们将发现，大多数时候言语并不能传达可信的信息，行胜于言。信息劣势方应当留意信息优势方的行为，而不是他的言语。而信息优势方也知道对手会这样解读自己的行为，反过来他也会为了利用其信息内容而控制自己的行为。

在进行策略互动博弈时，你可能会拥有（对自己）**有利**（good）的信息，

即，若对手获悉该信息，则他们就会采取有利于增加你的支付的行动。或者你可能会拥有**不利**（bad）信息，这类信息一旦披露则会导致对手采取不利于你的行动。所以，你会考量并采取行动，诱导对手认为你的信息是（对自己）“有利”的。这种行动就是所谓的**信号**（signals），使用信号的策略称为**信号传递**（signaling）。反过来，若对手倾向于认为你的信息是对自己“不利”的，你也许可以通过混淆其视听来阻止他们做出这一推断。这个策略，叫做**信号干扰**（signal jamming），它是一种典型的混合策略，因为混合策略的随机性使得对手的推断是不准确的。

如果对手掌握的信息比你多，或是他们的行动不会被你直接察觉，你可以采取策略以减少信息劣势。让对手披露其信息的策略称为**甄别**（screening），用于甄别对手的目的的方法叫做**甄别机制**（screening device）。[2]通过基于对手行动的可观察结果对其进行奖励或惩罚，试图影响对手的不可观察行动的策略，即所谓的**激励方案**（incentive scheme）。我们会在本章第3节到第5节详细说明这些概念。

9.2 直接沟通或“廉价磋商”

前面指出，关于策略博弈中涉及私人信息的问题，直接沟通是最简单的方式，但可能不具可信性。若说谎可以增加其支付，其他的参与人当然会选择说谎。不过，有一种博弈理论家称为**廉价磋商**（cheap talk）的直接沟通方式，可用于某些情况以帮助参与人从多个可能的均衡中进行选择。这样的博弈中，使用直接沟通达成的均衡称为**廉价磋商均衡**（cheap talk equilibrium）。

如果参与人的利益一致，那么直接的信息沟通将非常有用。第4章的安全博弈提供了此类情况最极端的例子。我们重新把该博弈的支付表（表4—12）画在下面（表9—1）。

表9—1　　安全博弈

		萨莉	
		星巴克	本地咖啡店
哈里	星巴克	1，1	0，0
	本地咖啡店	0，0	2，2

这个博弈中，哈里和萨莉的利益是完全一致的，他们都想见面，并且偏好于在本地咖啡店见面。问题是博弈是在无协调方式之下进行的，他们都在不知道对方之选择的情况下独立地做出自己的选择。然而，假如哈里在双方决定前有机会给萨莉发条短消息（或者萨莉有机会问一下哈里），短消息内容是：“我正去本地咖啡店。”萨莉没有理由认为哈里在说谎。如果萨莉相信哈里，萨莉就应选择本地咖啡店；并且，若哈里认为萨莉相信他，从而使得自己（哈里）的信息可信，其最好的选择也是去本地咖啡店。所以直接沟通很容易达成相互偏好的结果。这

就是为什么我们在第 4 章考虑此博弈时必须假定不能进行此类沟通的真正原因。不妨回忆一下两人来自不同的班级且直到见面前最后一刻也不知晓对方电话的情况。

让我们用更博弈理论化的术语考察一下安全博弈中允许直接沟通的结果。我们构造了一个两阶段博弈，在第一阶段只有哈里行动，其行动就是发短消息给萨莉。在第二阶段，最初的同时行动博弈开始进行。在整个两阶段博弈中，我们可得到逆推均衡，均衡策略（完整的行动计划）如下：第二阶段每个人的行动计划是“若哈里在第一阶段的消息是‘我去星巴克’则我也去星巴克；若哈里在第一阶段的消息是‘我去本地咖啡店’则我也选择本地咖啡店。”（记住，序贯博弈中参与人必须制定出详细完整的行动计划。）第一阶段哈里的行动是发送消息“我去本地咖啡店”。要确认这的确是一个两阶段博弈的逆推均衡很容易，我们把它留给读者自己去完成。

然而，这个廉价磋商产生了作用的均衡并非唯一的逆推均衡。考虑如下策略：第二阶段每个参与人的行动计划是，无论哈里第一阶段的信息是什么，两人都选择去星巴克；而哈里第一阶段的信息可以是任意的。我们可以确认，这也是逆推均衡。无论哈里第一阶段的信息是什么，若一方去星巴克，则对另一方来说其最优选择也是去星巴克。所以，在第二阶段的子博弈中，每个子博弈——哈里发送两条消息中任意一条之后的博弈——都可以使选择星巴克成为子博弈纳什均衡。从而，在第一阶段，哈里也知道其消息将毫无意义，无论选择发哪个消息都是无差异的。

哈里之消息并非无关紧要的廉价磋商均衡中，可产生更高的支付，我们很自然地会想到它将被选定为一个聚点（focal point）。但是，历史和文化因素也会使其他均衡有可能出现。例如，由于博弈之外的某些原因，哈里落下了完全不可信赖的名声，比如他可能是个老油条或神经兮兮的人。明知道哈里经常如此，人们对他的说法就会充耳不闻，萨莉也可能把它当耳边风。

这样的问题存在于所有的沟通博弈之中。沟通意见被忽略时，常常会产生其他与沟通意见无关的均衡。博弈理论家称之为**废话均衡**（babbling equilibrium）。然而，知道了废话均衡存在以后，我们将把注意力放到廉价磋商均衡上来，这样的均衡中沟通还是有一些作用的。

直接沟通的可信性取决于博弈双方利益一致的程度。安全博弈的一个深刻反例是参与人利益完全冲突的博弈，即零和博弈。表 4—15 的网球博弈是一个很好的例子。我们把它的支付表重新画在下表 9—2 中。请记住表中的支付是埃弗特的成功率；也请记住该博弈只有一个混合策略纳什均衡（在第 7 章推导过），均衡状态下埃弗特的期望支付是 62。

表 9—2　　　　网球博弈

		纳芙拉蒂洛娃	
		边线球	对角球
埃弗特	边线球	50	80
	对角球	90	20

现在我们考虑构造一个两阶段博弈。在第一阶段，埃弗特有机会发送信息给纳芙拉蒂洛娃；在第二阶段，则进行表 9—2 的同时行动博弈。何为逆推均衡？

很明显，纳芙拉蒂洛娃不会相信埃弗特的任何消息。例如，若埃弗特的消息是“我发底线球（DL）”，而纳芙拉蒂洛娃相信了她，则纳芙拉蒂洛娃将防守底线球；但是，若埃弗特认为纳芙拉蒂洛娃会这样做，则埃弗特的最佳选择就是发斜线球（CC）；从第二个层面思考，纳芙拉蒂洛娃应看穿这一点，并不会相信对方发底线球的信息。

但是问题没这么简单。纳芙拉蒂洛娃不应认为埃弗特一定会朝着其所声称的相反方向发球。假如埃弗特的信息是“我发底线球”，而纳芙拉蒂洛娃想“她是在诱我上当，因此我应判定她将发斜线球”，这将导致纳芙拉蒂洛娃防守斜线球。但是若埃弗特认为纳芙拉蒂洛娃以这种幼稚的方式不相信她，那么埃弗特就最终选择发斜线球。而纳芙拉蒂洛娃也应当看穿这一点。[3]

所以，纳芙拉蒂洛娃不相信埃弗特应该意味着她毫不理会埃弗特的信息。从而，两阶段博弈就只有废话均衡。第二阶段两人的行动与原始的均衡无异，埃弗特第一阶段的信息可以是任意的。以上结论对所有的零和博弈都成立。

但是，在更一般的既有冲突又有共同利益的博弈中，情况会如何？此类博弈中参与人利益仅部分一致，直接沟通是否可信，取决于冲突与合作的部分是如何组合在一起的。所以，我们对此类博弈中廉价磋商均衡和废话均衡都考察一下。

考虑一个多重均衡的博弈，一方偏好某个均衡，另一方偏好另一个均衡，但是任一个均衡都比其他非均衡结果好。例如性别战，我们将其支付表（表4—13）重画在表 9—3 中。

表 9—3　　性别战

		萨莉	
		星巴克	本地咖啡店
哈里	星巴克	2，1	0，0
	本地咖啡店	0，0	1，2

假设哈里有机会发送消息，则两阶段博弈有一个逆推均衡，哈里在第一阶段发送消息“我去星巴克”，在第二阶段两人的行动计划是选择他消息中所说的地点。这里，存在一个廉价磋商均衡，并且获得发送消息的机会可使参与人选择自己偏爱的结果。[4]

支付部分一致博弈的另一个例子来自你可能经历过的一种情形，你如果没经历过也会在你开始赚钱或投资时很快遇到。当股票经纪人建议你买某一支股票时，他既可能是以此作为与你发展长期关系的一部分，以便为你做长期的商业代理获取好处；也可能是鼓吹他持有的某支垃圾股票以便脱手获利。你必须猜测其支付中两种可能的相对重要性。

设想刚好存在两种可能性：股票具有涨的前景和跌的前景。若是前者，你应当购买；若是后者，则应当抛售，或卖空（sell short）。表 9—4 表示了给定涨、跌两种可能性以及买、卖两种行动，你在各种情况下的支付。注意，这不是一个

博弈的支付矩阵；表中的列并非策略参与人的选择。

表 9—4　　你从投资决策中获得的支付

		股票	
		涨	跌
你	买	1	−1
	卖	−1	1

经纪人知道股票实际上是涨是跌，而你却不知道。他可以建议你买或卖，你应言听计从吗？问题的答案取决于你的经纪人在同样情况下的支付。假设经纪人的支付是两种考量的混合，一种考量是与你建立长期关系，这一部分支付来自你的支付。但是若他急于脱手而说服你购买一支糟糕的股票，则他也可得到额外的回扣 X。其支付表示在表 9—5 中。

表 9—5　　你的经纪人从你的投资决策中获得的支付

		股票	
		涨	跌
你	买	1	$-1+X$
	卖	−1	1

这里存在诚实沟通下的廉价磋商均衡吗？本例中，经纪人在博弈第一阶段向你发送信息，其信息可以是买或卖，这取决于他观察到股票是涨还是跌。在第二阶段，你决定是买还是卖，这取决于经纪人发给你的信息。这样，我们需要寻找到一个均衡，该均衡中经纪人的策略是诚实的（涨的时候说买，跌就说卖），而你的策略是听从其建议。

给定经纪人发送的是诚实信息，显然对于你而言最佳选择是听从建议。给定你听从建议，经纪人的策略又会如何？假如他知道股票是涨的，若他发出买的信息，你将买进股票而他会获利 1；若他说卖，你会卖出股票而他的支付是 −1。所以，其策略中“若股票涨就说买”的部分对他来说的确是最优的。现在，再假设他知道股票是跌的，若他说卖，你会卖掉股票而他得到 1；若他说买，你会买进而他将获得 $-1+X$。所以对于他而言只有在 $1>-1+X$（或 $X<2$）的情况下发送诚实消息才是最优的。因此，只要经纪人从你所买的垃圾股票中获取的利益并不“太大”，则与自经纪人的直接沟通是可信的，该博弈存在廉价磋商均衡。

然而，若 $X>2$，且你的策略是听从其建议，则经纪人的最优反应是说买，不管股票涨还是跌。但是，如果他这样做，那么听从其建议就不再是你的最优策略。你必须对其消息不置可否并重回到你对股票涨跌的先验估计。这种情况下，只有废话均衡是可能的。

在这些例子中，可获得的消息是简单的二元信息——去星巴克或去本地咖啡店，以及买进股票或卖出股票。若信息更丰富，情况又如何？例如，假如经纪人可以发送给你一个数字 g 代表他对股票价格增长率的估计，并且数字可以在一个

连续区间内变动。那么，只要经纪人可从你购买垃圾股票的行为中获得额外利益，他就会过高地报告 g。因此完全诚实地沟通就不可能。但部分地诚实披露是有可能的。即，增长率连续变动的范围应当细分成组——如 0%到 1%，1%到 2%，依此类推，这样经纪人会发现，增长率落在某些区间时，对你讲实话对于他而言是最优的，而在此基础上你也发现接受其建议对于你来说也是最优的。但是，我们必须将这一思想的进一步解释留给更高级的分析。[5]

9.3 信号传递和甄别

当直接沟通不能产生廉价磋商均衡时，则行动可胜于言语。原因是，说来容易做来难——口水太廉价——而行动则有其直接代价，关乎其支付。如果行动的代价以恰当的方式依赖于私有信息本身的一些内容，则这些行动就是隐匿信息的可信证据。我们现在来仔细讨论这一思想。

大多数学生都希望毕业之后能成为美国精英员工之一，而公司老板希望他们具备应有的素质和技能——努力工作、数字计算、逻辑等能力。你比老板更清楚自己的能力。他可以跟你面谈或让你参加考试，但是因为时间和资源的关系，他通过这些方法能获得的信息是有限的。你仅仅是口头宣称自己具备应有的素质是难以令人信服的，还得需要客观证据。

老板可以找到什么证据？你又能提供什么证据？回忆本章第 1 节，我们把前者称为“甄别机制”，老板用它们来确认你的素质和技能，即甄别。你主动使用的证据叫信号，当你使用它们时，你就在进行信号传递。有时，类似甚至相同的方法既可用于信号传递也可用于甄别。

在这个例子中，如果你在大学中选修了难度高且任务重的课程（并通过了考试），则你对课程的选择通常可作为你素质能力和分析技能的证据。现在让我们认识一下课程选择作为甄别机制的作用。

为简化起见，假设在公司老板眼里大学毕业生只有两种类型：A 类（能力强）和 C 类（能力不足）。老板愿意给 A 类员工 150 000 美元年薪，给 C 类 100 000美元年薪。假设有很多潜在的老板在争夺数量有限的求职者，所以老板们必须支付其愿意支付工资的最高金额。既然老板不能直接观察求职者的类型，他就必须寻找可靠的手段来识别。

假设两类毕业生的差别是 A 类能在大学中接受难度高任务重的课程，但 C 类接受这些课程比较吃力。为了修一门高难度课程，他们都需要牺牲课余活动，但 A 类的牺牲会较少，C 类的牺牲会较大，或 A 类比 C 类更能忍受沉重的压力。假设 A 类把修此类课程的代价视为一年 6 000 美元的薪水，而 C 视为一年 9 000 美元，试问老板可以用这种差异来分辨 A 类与 C 类吗？

考虑如下策略：任何修过 n 门高难度课程的毕业生将被视为 A 类，年薪 150 000美元；而少于 n 门者则会被视为 C 类，年薪是 100 000 美元。该策略的目标是要让 A 类有修高难度课程的动机，而 C 类没有这个想法。当然，没有人喜

欢多修高难度课程，所以他们的选择是要么刚好修 n 门以符合 A 类的资格，或者干脆不修任何高难度课程而被视为 C 类——他们念大学仅仅是为了混张文凭粉饰一下脸面而已。

老板用于识别 A 类跟 C 类的标准——也就是高难度课程的修课门数——应该足够严格到让 C 类无意达到标准，但又不能太严格而让 A 类放弃修课。正确的 n 值必须让 C 类暴露自己的身份，而不是凭借修课来模仿 A 类的行为，所以我们要求[6]

$$100\ 000 \geqslant 150\ 000-9\ 000n\text{，或 } 9n \geqslant 50\text{，或 } n \geqslant 5.56$$

同样，让 A 类乐于修 n 门高难度课程的条件是

$$150\ 000-6\ 000n \geqslant 100\ 000\text{，或 } 6n \leqslant 50\text{，或 } n \leqslant 8.33$$

这些约束条件使求职者的动机和老板的愿望保持了一致，或者说，它们使得求职者通过其行动披露技能的真实信息是最优的。所以这些条件被称作该问题的**激励相容约束**（incentive-compatibility constraint）。由于课程门数应是整数，所以满足条件的 n 必须是 6，7 或 8。[7]

要满足这些条件需要让两种类型学生选修高难度课程的成本之间存在差异：对于能力强的学生来说，付出的代价必须足够小，如此老板才能辨认。当达到约束条件时，老板的策略就可以令两种类型的学生做出不同反应，因而暴露他们的类型。这称作基于**自选择**（self-selection）的**类型分离**（separation of types）。

这里我们不假设高难度课程可以提供额外的技能或工作习惯，让 C 类转变成 A 类。在我们的设想中，高难度课程仅仅是用于分辨不同类型个人的工具。换句话说，它们只具备甄别功能。

事实上，教育的确可以增强生产能力，但它也具有这里所说的甄别功能和信号传递功能。在本例中，我们发现教育可能仅仅承担了后一功能；而现实中，进一步教育深造的相应结果已远非额外的生产能力所能解释，同时进一步的教育也带来了新的成本——信息不对称的成本。

当我们用必要的高难度课程修课门数来进行甄别时，A 类就承担了代价。假若只用最小修课门数（即 $n=6$）来筛选，则 A 类付出的货币代价是 $6\times 6\ 000=36\ 000$美元。这就是信息不对称的代价。若个人的类型可以直接观察或识别，这种代价是不会存在的。或者整个社会都是由 A 类人口组成，这种代价也不会存在。A 类之所以付出此类代价是因为存在一些 C 类人口，而他们（或其老板）必须设法做出识别。[8]

若不采取分离策略而 A 类也不承担代价，情况会好些吗？分离政策下，A 类得到年薪 150 000 美元但是同时付出货币代价 36 000 美元，所以其货币等价支付为 114 000 美元。而 C 类只得到年薪 100 000 美元。如果不把他们分离开来会怎样？

如果老板不采取甄别机制，他们就不得不对所有求职者一视同仁，支付相同的薪水。这被称作**类型混同**（pooling of types）。[9]在竞争的劳动力市场上，共同薪水（common salary）会等于各种不同类型人口的平均薪酬水平，而这取决于不同类型在人口中的比例。

例如，若人口的20%属于A类，80%属于C类，则混同下的共同薪水将是

$$0.2\times150\ 000+0.8\times100\ 000=110\ 000 \text{ 美元}$$

所以A类将偏好分离的情形，因为分离状态下可得到114 000美元而不是混同状态下的110 000美元。但是，若两种类型的人口比例各占50%，则混同下的共同薪水将是125 000美元，那么A类在分离状态的处境将比在混同状态下糟糕。C类则总是偏好混同均衡，因为A类的存在使得混同下的共同薪水总会高于分离时C类的薪水100 000美元。

然而，在大量雇主和雇员使用甄别和信号传递进行竞争的时候，即使两类求职者都偏好混同结果，它也不可能是均衡。比如，假设两类人口比例各占50%，初始均衡是类型混同下的薪水125 000美元，则某个老板可以宣称他将付132 000美元给修过一门高难度课程的人，A类觉得划算是因为修课只花6 000美元而薪水却增加了7 000美元；C类会觉得不划算，因为修课花9 000美元，超过了增加的薪水7 000美元。某个老板用这招来吸引A类，而且只花132 000就可雇用到价值150 000美元的人，所以他的类型分离策略大获收益。

但是，这个老板对类型混同薪水的背离牵一发而动全局，引发了竞争性雇主之间的调整过程，并导致原先的混同情况被瓦解。由于许多A类员工都被吸引到这个老板那里去了，其他老板雇用的员工平均能力就会降低，最后他们不能负担125 000美元的薪水，于是类型混同下的薪水将降低。而随着混同下的薪水的降低，跟某老板提出的分离下的薪水132 000美元的差距扩大，C类也会觉得去修一门高难度课程是值得的。此时，这个老板就必须把标准提高到修两门课，让C类再次觉得修课是不划算的。其他老板也会用同样的政策，如果他们也想吸引A类员工的话。过程一直持续，到达我们先前所说的分离均衡为止。

即使老板们不用这些手段来吸引A类，混同状态的薪水为125 000美元的A也会去修一门高难度课程，然后拿着成绩单去跟老板说："我的成绩单上有一门高难度课程，这应该能证明我是A类员工，C类员工是不能对你如此保证的。所以我现在要求132 000美元的薪水。"给定事实如此，其主张是有效的，老板也会觉得答应他是符合自己的利益的，因为A类员工有150 000美元的价值，但只领132 000美元的薪水。其他的A类员工也可以这么做。然后像刚才一样，这过程又会一直持续，直到达到分离均衡。唯一的差别在于由谁发起，现在是由A类员工发起，他选择接受额外的教育以证明其类型；这是信号传递而不是甄别。

虽然混同结果对所有人有利，但追求个人利益却会导致分离均衡。这就像有大量参与人的囚徒困境博弈，所以人们不可避免地要为信息不对称付出代价。

在参与人缺乏直接竞争的其他情形中，混同均衡与分离均衡都可能出现，但这还不是全部的可能结果。在完全分离均衡中，结果可完全披露参与人的类型；在混同均衡中，结果不能提供任何有关类型的信息。有些博弈可以得到中立均衡（intermediate equilibrium）或半分离均衡（semiseparating equilibrium），在这样的均衡中，结果可以部分地提供有关参与人类型的信息。分析这类博弈要求稍艰深一点的概念和技术，所以我们直到第5节才会给出例子。

到此为止我们只考虑了两种类型，但是其思想可很快推广到多种类型。假设

有好几种类型：A，B，C，……依能力强弱排列，则可以建立一连串不同等级的标准，能力最弱的人选择不受教育，比能力最弱者好一点的人选择接受一点教育，依此类推，所有人都会自选择可识别出其类型的标准。

在这个例子中，我们发展出的思想是，老板可以通过毕业生修课的难易程度来确认他的能力，这就是甄别机制的例子。但时不时我们也谈到，学生也可以主动采取同样的行动作为信号。正如先前所说，信号传递和甄别事实上是结伴而行的，尽管有些情况下均衡可能不同，这取决于博弈过程中谁采取主动，以及博弈各方之间的竞争环境。我们把这些细节留给更高级的博弈理论去分析，在此我们仅对信号传递做一点深入说明。

假设你是信息优势方，并且有一个令人信服的行动可以传递出对自己有利的信息。如果你无法传递出这个信号，其他人就会持有于你不利的信息假设。这样看来，信号传递就像在比赛胆量——如果你拒绝出招，那你就已经出招并且输了。

当你在选择课程成绩是采用字母计分方式[①]还是采用“合格—不合格”二元计分方式时，你应当记住前面的道理。该课程学生的成绩会广泛分布在各个成绩等级上。假定平均成绩是 B，则学生就会明白自己成绩的好坏。那些有信心拿到 A＋的学生有强烈的动机要求采取字母计分方式。如果他们这样做，其他学生的平均成绩就会低于 B，比如变成 B－，因为 A＋的人已经占去了等级高的部分。而在剩下的学生中，那些有望得到 A 的学生有动机去另选别的课程，这又降低了该课程其他学生的平均成绩。如此下去，最后有望获得 C 和 D 的学生只能选择采取二元计分（合格或不合格）的课程了。这正是那些阅读成绩单的人（公司老板或研究生招生委员会）如此看重计分方式的原因。

9.4 激励报酬

假如你是一家公司的老板，正在物色一名经理去负责一个项目。项目的结果是不确定的，成功的概率取决于经理的努力。如果该项目成功，则可赚回600 000美元。若经理只是一般努力则成功率为 60％（0.6），若经理很努力则会提高到 80％。

若经理很努力则会令其付出主观代价。例如，他必须为项目殚精竭虑而疏忽家庭。若他只是一般努力，你给他 100 000 美元就可以了；若他很努力，则你还要额外给他 50 000 美元。

问题是值不值得付给他额外的薪水呢？若不付，你只有 0.6 的概率获得600 000美元（期望值为 360 000 美元），支付经理薪酬100 000美元，所以期望的净利润为 260 000 美元。如果支付，期望利润为 0.8 × 600 000 － 150 000 ＝480 000－150 000＝330 000 美元。因此答案是应当支付；向经理额外支付 50 000美元让他更努力，对于你来说也更有利可图。

① 即以 A，B，C，D 评定成绩。——译者注

但如何实施这样的安排？给经理的合约中可以写明其基本薪酬为 100 000 美元，如果他很努力就会得到额外的 50 000 美元。但怎知他是否努力？经理可能只有一般程度的努力但还是把额外的 50 000 美元拿走，因为他可以宣称他都把公事带回家，趁晚上或周末做，而你却无法验证。如果项目失败了，他可以将失败归结为运气不好，毕竟，即使很努力也还是有 1/5 的可能性会失败。而一旦你与经理就其努力程度发生争执，告上法院，法官又如何判决？

如果努力程度不可观察，也无法在法庭得以证实以避免争端，则经理的合约必须基于某些可观察和可验证的东西。在目前这个例子中，那就是项目的成败。因为成败的概率与努力程度相关，它可以告诉我们关于努力程度的信息，虽不完整，但还是可以用来激发努力。

考虑一个由基本薪酬 s 和奖金 b 组成的报酬结构，b 只有在项目成功时才支付。因此，若经理只是一般努力，其期望收入为 $s+0.6b$；若他很努力，则其期望收入为 $s+0.8b$。所以经理因额外的努力而获得的期望额外收入为 $(s+0.8b)-(s+0.6b)=0.2b$，要让他付出额外努力，下列不等式必须成立：

$$0.2b \geqslant 50\,000 \quad 或 \quad b \geqslant 250\,000$$

其解释很简单：奖金乘以额外努力导致的成功概率增加量，等于经理额外努力的期望额外收入；这笔期望额外收入不应低于经理付出额外努力的代价。

所以，足够高的奖金可以对经理的额外努力产生足够大的激励。同第 3 节一样，这被称为报酬结构的激励兼容条件或激励相容约束。

还有另一种条件，也就是整个报酬结构必须好到让经理愿意为你工作。当达到激励兼容条件时，经理会为你更加卖命地工作，他的预期收入是 $(s+0.8b)$，但他也至少需要 150 000 美元才值得为你付出，所以你提供的报酬必须满足

$$s+0.8b \geqslant 150\,000$$

这称为**参与条件**（participation condition）或**约束**（constraint）。

而你试图使你的利益达到最大化，所以你希望在满足激励相容约束和参与约束下把经理的报酬定在尽可能低的水平，即 150 000 美元，这意味着要选择满足参与条件的最小的 s，或选择 $s=150\,000-0.8b$；但是，既然 b 必须不小于 250 000，则 s 就不会超过 $150\,000-0.8\times250\,000=150\,000-200\,000=-50\,000$。

负的基本薪酬表示什么意思？对此有两种可能的解释：其一，负的基本薪酬可能表示经理在该项目中必须做的投资，奖金就可以解释为股权利益或合伙收益；其二，负的基本薪酬也可解释为项目失败的惩罚。

有时候，这些可能的办法都无效。有资格胜任经理职位的人可能没有钱投资，而法律则禁止对员工罚款。这样的情况下，总薪酬就不能降低到刚符合参与条件。例如，假设基本薪酬 s 必须是非负的，即便是奖金水平达到满足激励相容约束的最低水平，即 $b=250\,000$，而你能给经理的最低基本薪酬是 0，则经理的预期总报酬是 $0+0.8\times250\,000=200\,000$ 美元。你被迫过度满足了参与约束。

如果努力程度可以被直接观察到并可验证，则你会签下这样的合约：付基本薪酬 100 000 美元，如果看到经理很努力，则额外给他50 000美元，总共 150 000 美元。如果努力程度不能被观察到且不可验证，你就需要付给他 200 000 美元，

以激励他更加努力。这额外的 50 000 美元是观察难题——也就是博弈中信息非对称——所造成的代价。这样的代价通常存在于非对称信息的情况下。博弈论告诉了我们如何对付非对称信息，但这需要付出代价，而且通常是信息劣势方来承担这个代价。

为了克服信息问题而付出额外的代价是否值得？在这个例子中，支付额外代价换来 0.8 的概率获得 600 000 美元，你的期望净利润为

$$0.8\times600\ 000-200\ 000=480\ 000-200\ 000=280\ 000 \text{ 美元}$$

你也可以只要求经理付出一般程度努力，所以只需要给他 100 000 美元的基本薪酬，换得 0.6 的成功率，期望净利润为

$$0.6\times600\ 000-100\ 000=360\ 000-100\ 000=260\ 000 \text{ 美元}$$

所以，即使要承担信息非对称引起的额外代价，使用激励方案仍可使你得到更多的期望利润。

不过，并非所有情况皆是如此。把这个例子中的数字改变一下就可以有不同的结果：如果项目成功仅赚回 400 000 美元而不是600 000美元，则你采取低薪酬，经理付出低努力，对你反而是更有利的。

这个问题的本质与保险公司遇到的问题类似。保险公司的客户有遭受重大损失的风险——比如遭遇盗窃或火灾，但客户是否采取良好的预防措施或者是否小心留意也可以影响损失发生的概率。如果参加了保险，则当你不能肯定在离家时是否反锁了大门时，可能就懒得回去检查一下。正如前面所说，保险业将此问题称为道德风险。有时还会有故意骗保，比如自己在家纵火去诈领保险金，不过很多情况不一定会牵扯到道德品质问题。

非对称信息情况和激励问题的经济分析已经接纳了这个术语。一般地，任何个人行为不可观察和验证的策略互动，都称作**道德风险**（moral hazard）。保险公司的例子中，通过自付额或共同保险让受保人承担部分风险，道德风险可以得到控制。这种做法与激励方案一样；事实上，它正是经理签约问题的翻版：经理因项目成功而获得奖金，但他的努力程度会影响项目成功的概率；受保人因为灾难而必须承担部分损失，但他的行为会影响损失发生的概率。

与我们的简单例子相比，真实世界的道德风险问题或多或少都更严重一些，对付道德风险的激励方案在多大程度上是可行的，要取决于现实环境。我们曾假设项目的成败可以客观地观察到，但实际上经理也许可以操纵结果使得项目看起来是成功的，而真正的结果是在他得到奖金离开之后再显示出来。这种可能性并不夸张：许多公司在 2000 年还在一片喝彩声中被看做新经济奇迹，到 2002 年却只剩下一副空壳；他们先前宣称的收入和回报不过是会计处理上的花招。所有者（外部股东）深受其害，而经理们却钱袋鼓鼓地离开了。

另一方面，也存在诸多办法可使所有者获得更为准确的有关经理努力程度的信息。比如，如果有几个经理并行运作类似的项目，并且每个人获得结果的运气因素相互正向关联，那么所有者就可以通过比较他们的结果来判断每个经理的努力程度。如果其他经理得到了好的结果，那么某个得到坏结果的经理就不能抱怨说自己运气太差。只有所有经理进行**共谋**（conspiracy）才能挫败这样的相对评

估，但是这样的共谋受制于经理的囚徒困境，即若其他经理偷懒，那么自己努力一点就可得到可观的奖励。

如果所有者和经理重复互动并且每一时期的幸运因素是独立的，也可以降低道德风险。尽管少量的偶然失败可以归因于运气太背，但一连串接踵而至的失败极不可能发生。因此，经理的长期偷懒行为将受到有力遏制，而更好的激励方案也可以设计出来。

9.5 信号传递博弈中的均衡

在本章第 3 节、第 4 节中，我们解释了甄别和信号传递的一般概念，以及运用这些策略后可能产生的分离和混同结果。我们也看到了在存在众多老板和员工的情形中（此时薪酬由市场竞争决定）如何运用这些概念。然而，我们尚未说明两位信息不同的参与人之间进行的博弈，并且尚未求解其均衡。现在，我们用一个例子来讲解此类情况。特别地，我们将看到分离或混同皆可成为均衡，并且一种**局部信息披露**（partially revealing）或**半分离均衡**（semiseparating equilibrium）也会在博弈中出现。不过有一些概念和分析会比较复杂，初学者可以忽略本节内容，这不会影响阅读的连续性。

我们下面要介绍的博弈有两个参与人：一个攻方，一个守方。守方或强或弱。攻方需要决定是否侵略，但他不知道守方的确切类型。攻方假定守方为强类型的概率是 0.25。[10]守方知道这是攻方的信念，并且攻方知道守方知道，依此类推。换句话说，这是共同知识。如果攻方侵略而守方其实是强类型，就会爆发一场战争，双方均付出 10 单位代价；反之，若守方其实是弱类型，不战而降，则守方损失 5 单位而攻方得到 5 单位。最后，守方有机会先发制人采取一个行动，即我们所说的信号，无论守方为何种类型该行动都将付出 6 单位代价。攻方观察到守方是否给出信号然后采取行动。我们的目的，是要确定强类型守方是否会采取行动传递可信的类型信号。博弈的结果依赖于我们对概率和支付数字的具体选择。我们将改变数字来说明结果如何变化，并在稍后给出一般性的代数说明。

这个例子抓住了现实中诸多博弈的特征。在国际关系中，一个国家对另一个国家虎视眈眈，却不知道对方的抵抗决心有多大。防守国可以通过公开战备来防止被侵略，但如果它确实是弱小的，则一旦侵略发生这些战备立马就灰飞烟灭。[11]在经济学中，某个企业也许在考虑是否进入已有一个垄断企业的行业，却不知道会与该垄断企业爆发多激烈的竞争。该垄断企业可以采取某些行动——建设发动价格战的生产能力，或者制定较低的价格作为其低成本生产的证据——使试图进入者相信一旦进入则意味着一场与强者对阵代价沉重的价格战，从而遏止其进入。[12]在自然界，有些动物有能力在面临威胁时暂时变得看起来更大更强壮，即使这种能力并不能在真正遭遇战斗时帮助它们。[13]在治安混乱的贫困街区，年轻人展示出他们的名贵衣服、珠宝以及其他“需要守护的物品”，这种“自我展示宣称：这个男人可以照顾好自己，不要去招惹他”[14]。

传递强硬信号的想法很吸引人但却是有问题的。罗伯特·杰维斯（Robert Jervis）在 20 世纪 70 年代早期就曾指出，弱小的防守方比强大的防守方更需要避免侵略，因此信号传递行动应该被解读为弱小而不是强大。不过，如果这是侵略者的解读方式，那为什么任何人都不辞辛劳地提供代价沉重的信号？为了理清这令人混淆的循环论断，就必须考察完整的博弈及其均衡。我们从图 9—1 的博弈树开始。

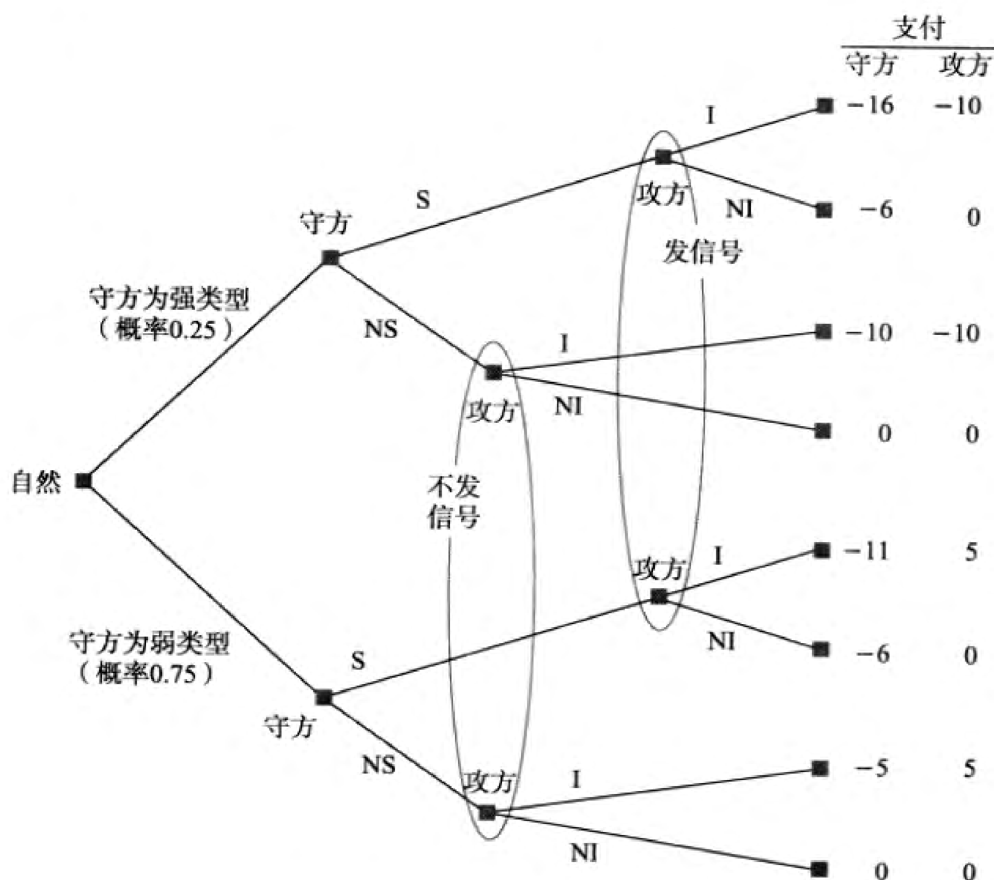


图 9—1　传递强硬信号博弈

当是其他因素而不是参与人的随机行为导致不确定性产生的时候，我们通常会引入一个中立参与人（neutral player），即自然（Nature），由它先以 0.25 的概率选择守方为强类型或以 0.75 的概率选择守方为弱类型，然后启动博弈。在每一个自然选择结果的结点上，守方知道其类型，并决定发送（S）或不发送（NS）信号。由此攻方面临四个结点，在每个结点上攻方需决定侵略（I）或不侵略（NI）。但是，攻方只观察到守方的行动而未能观察到自然的行动，所以他无法辨别出发送信号的守方之强弱类型，也无法辨别出不发送信号的守方之强弱类型。因此攻方有两个信息集：一个信息集由守方选择 S 的两个结点组成，另一个由守方选择 NS 的两个结点组成。在每个信息集的两个结点上，攻方必须选择同样的行动 I 或者 NI。

这个博弈中正确的均衡概念是什么？序贯行动博弈逆推求解的思想大部分还可派上用场。攻方在每一个信息集的选择对于他应当是最优的，而每种类型的守

方在考虑到攻方的后续选择时所进行的决策对于其自身也应是最优的。攻方观察到守方的行动 S 或 NS 后，便可从中进行推断。例如，若均衡策略要求守方必须在强类型时选择 S 而弱类型时选择 NS，那么观察到守方选择 S 就足可证明守方为强类型。也就是说，一旦观察到守方选择 S，攻方就确切地知道守方是强类型；强类型的概率变为 1，而不再是缺少进一步信息下的先验概率 0.25。要在这种类型的博弈中找出均衡，我们应当假设攻方是理性的，他在选择是否侵略时将做出所有这样的推断；同时也需要假设守方是理性的，他在选择是否发送信号时，会意识到攻方将会做出这些推断。

在概率论中，这种基于临时观察的概率修正是由贝叶斯法则（Bayes' rule）来确定的。[15]攻方遭遇两种类型的概率一旦修正就可能会影响其最优选择。所以，我们必须讨论逆推求解或均衡的子博弈完美（subgame-perfectness）概念，并要求：

（1）在每一个信息集，参与人按照可获得的信息（包括信息不明确时各个结点出现的概率）采取其最优行动；

（2）参与人从他们的观察中进行正确的推断，按照贝叶斯法则进行概率修正。

满足上述标准的结果，就是所谓的**完美贝叶斯均衡**（perfect Bayesian equilibrium，简记为 PBE）。

在前面的章节，我们是通过从终点向起点逆推的方式来寻找序贯行动博弈的均衡的。然而，在我们眼前这个博弈中，后行动者的行动却要求根据先行动者的行动来修正其信息。这导致了一个不可避免的循环，简单的单一逆向推理在这里是无效的。相反，我们首先必须针对所有参与人在所有阶段的选择制定一套完整的策略（完备的行动计划）并对信念形成（概率估计的修正）加以明确规定。在制定出策略并规定了信念形成过程之后，我们必须确定它们是否满足 PBE 的所有要求。下面我们用例子来讨论可能产生的不同结果，并检验 PBE 的存在性。

9.5.1 分离均衡

要得到分离均衡，守方的行动必须能完全揭示其类型，从而攻方可以确切地知道其所处的决策结点的位置，即他知道自己进行下一阶段博弈的起点在何处。我们将考虑守方采取行动诚实披露其类型的情况。[16]具体地，假设守方若为强类型则选择 S，若为弱类型就选择 NS。攻方通过观察到的行动推断守方的类型（观察到 S 则推断为强类型，观察到 NS 则推断为弱类型），并且在给定守方选择 NS 时自己选择 I，守方选择 S 时自己选择 NI。

这组策略在博弈中是均衡吗？要回答这个问题，我们需要检查其行动和信念是否满足先前所说的两个标准。首先，如果攻方发现自己在 S 信息集中（观察到 S），那么给定守方选择 S 的策略，当且仅当其为强类型时，攻方可以正确推断出他的确处于上方的信息结点（以 1 的概率面对强类型守方）；攻方在此结点的最佳选择是 NI，这使得其支付为 0，而选择 I 则支付为－10。同样，在 NS 信息集，攻方将推断出他的确处于下方的信息结点（以概率 1 面对弱类型守方），此时攻

方选择 I 可获得支付 5，选择 NI 获得 0。然后，我们再看两种类型的守方的最佳行动。考虑强类型守方，若他选择 S，则攻方相信他是强类型并选择 NI，因此守方将得到－6；如果他选择 NS，则攻方相信他是弱类型并选择 I，于是这个事实上为强类型的守方得到支付－10。因此这种情况下，一个强类型的守方将宁愿选择 S。采用同样的推理步骤我们还可得到，一个弱类型的守方，若选择 S 则得到－6（攻方选择 NI），若选择 NS 则得到－5，所以他宁愿选择行动 NS。

我们的分析表明，规定的策略和攻方的信念形成过程的确构成一个博弈均衡。每个参与人规定的行动对其自身都是最优的，而且从行为观察中获得了正确推断。所以，在给定的支付结构下，存在着守方诚实发送其类型信号的完全分离均衡。

但是，若对支付结构稍加改变，情况又会如何？假设传递信号的成本降低到 4，比（弱类型守方）不战而降的代价还要小。于是博弈改变成图 9—2 所示的博弈，而规定的策略已不再是均衡。攻方从观察到 S 推断守方为强类型以及从观察到 NS 推断守方为弱类型，仍旧在看到 S 时选择 NI，看到 NS 时选择 I。而强类型的守方仍宁愿选择 S（支付－4）而不是 NS（支付－10）。但是，弱类型的守方若选择 S 则支付为－4，若选择 NS 则支付为－5。弱类型守方的最佳选择不是 NS，PBE 的要求不再满足。显然，信号传递博弈中分离均衡（事实上任何均衡）的存在与否关键取决于与不同策略相联系的支付之大小。我们在 9.5.4 节将详细研讨此结论的一般意义，不过现在还是先仔细看看另外两种可能的均衡。

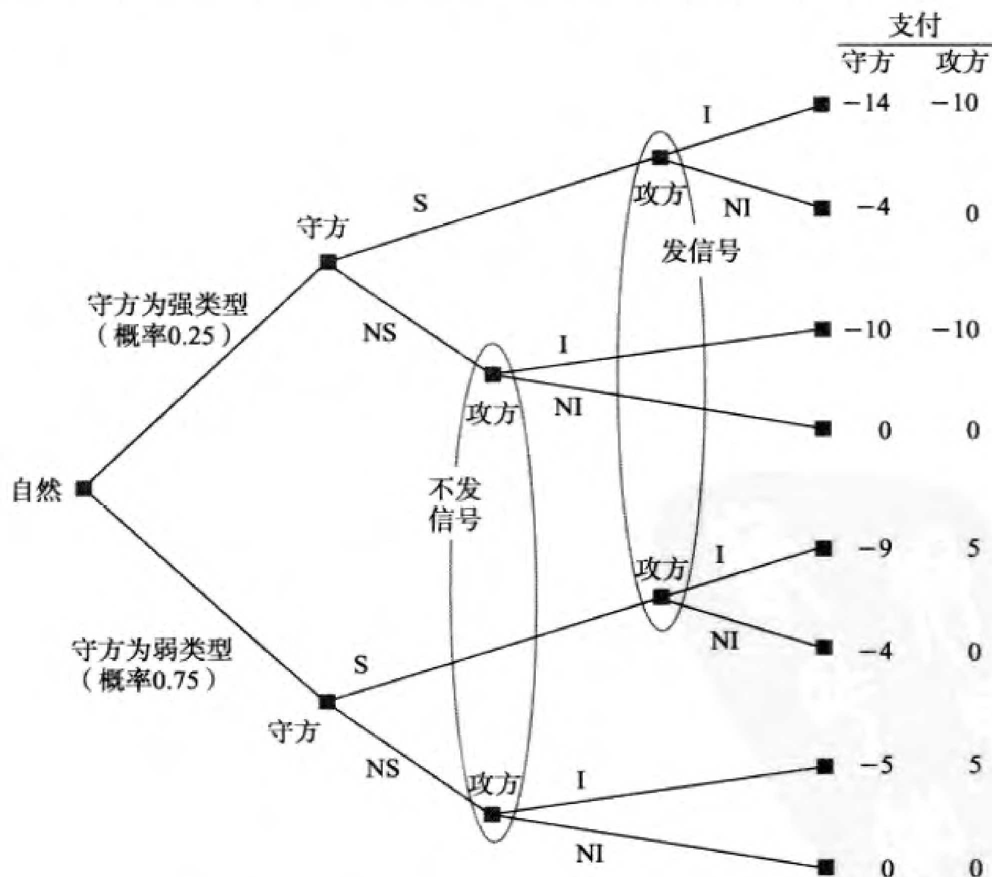


图 9—2　传递强硬信号博弈（信号成本降低的情况）

9.5.2 混同均衡

在上一小节末，我们看到支付改变可能导致分离均衡消失无踪，因为弱类型守方也宁愿选择 S。所以这种情况下，很可能存在一种均衡，在这一均衡中两种类型的守方选择同样的行动，即，类型是混同的。让我们考虑所有守方均选择 NS 的混同均衡。攻方观察到 NS 时，他无法修正自己的信息，只能根据已知其对手属强类型的概率为 0.25 来做出最优选择。或者，攻方观察到 S 时，我们假设他认为这种误发信息的概率对两种类型守方是一样的，也就是说即便观察到 S，攻方也不能修正其信息，从而其最优选择是在类型同等可能的假设下做出的。[17]

如果这组策略和信息修正过程是一个均衡，那么给定攻方的最优选择，它必须对两类守方选择 NS 都是最优的。攻方观察到 S 或 NS 并选择 I，若守方为强类型则攻方支付－10，若守方为弱类型则攻方支付 5；所以攻方选择 I 的期望支付为(－10×0.25)＋(5×0.75)＝1.25。同样，若攻方观察到 S 或 NS 并选择 NI，无论守方是何类型攻方都得到 0，因此他选择 NI 的期望支付是 0。所以，无论攻方观察到 S 还是 NS，他都宁愿选择 I。给定他的选择，强类型守方选择 S 则支付－14，选择 NS 则支付－10；弱类型守方选择 S 则支付－9，选择 NS 则支付－5。两类守方都宁愿选择 NS，所以我们规定的行动满足 PBE 的条件。

这个混同均衡的达成，部分原因是假设了攻方不能在观察到 S（均衡之外的行动）时获得信息。假若我们反过来提出一个不同的结果，其中两类守方均继续选择 NS，但是攻方在观察到 S 时可以修正其信息。此时，S 被解释为强类型的确切信号，而不是像刚才那样被解释成两类守方都会犯的失误。那么，当攻方观察到 S 时，他会选择 NI 得到支付 0，因为选择 I 就是与强者争锋只会得到支付为－10。当攻方观察到 NS，他像以前一样选择 I。明确攻方的选择后，弱类型守方会发现选择 S 将得到－4，选择 NS 会得到－5，所以他宁愿选 S；同样，强类型守方选 S 得到－4，选 NS 得到－10，所以他也宁愿选 S。但两类守方都选 NS 在这里并不是均衡，因为攻方将信号 S 解读为确切的强类型，但两类守方却因此而皆想传递信号 S。这个例子表明，在描述混同均衡时规定攻方信息修正过程很重要。

9.5.3 半分离均衡

在前面的例子中，给定具体支付和信息修正过程，则我们的博弈可能既不存在分离均衡也不存在混同均衡。然而，在这个信息非对称的博弈中，还有最后一种可能的均衡类型，即使得攻方和某类守方必须采取混合策略的均衡。这种结果可能在如下情形中发生：如果某一类型守方在攻方采取混合策略时无论选 S 还是 NS 其支付都一样，因此他也乐于混合其策略；但是另一类型的守方却偏好 S 和 NS 中的一个。例如，考虑所有强类型守方选择 S 而弱类型守方采取混合策略。

观察到 S 可以给攻方有关守方类型的某些信息——但不是全部信息，所以我们用半分离均衡来描述这种情况。攻方必须使用贝叶斯法则仔细修正其概率，并且均衡将要求他在观察到 S 时采取混合行动以使得弱类型守方在 S 和 NS 之间无差异。[18]

要确定所说的一组策略是否是博弈的一个均衡，就需要求解强弱两种类型守方和攻方所使用的混合概率。假设均衡中弱类型守方选择 S 的概率是 x，则 x 必须使得攻方选择 I 和 NI 无差异。但是攻方现在可以使用观察到的 S 修正其关于守方类型的概率判断。给定所有强类型守方将选择 S 而弱类型守方以概率 x 选择 S，则攻方预见到 S 被采取的概率就是 $0.25+0.75x$，这一结果来自第 7 章附录中的概率组合律：强类型守方的概率是 0.25 而其选择 S 的概率为 1，弱类型守方的概率是 0.75 而其选择 S 的概率为 x。既然观察到了 S，则攻方可因此修正关于守方属于强类型的概率。在出现 S 的概率中，归属于强类型守方的概率部分是 0.25；所以一旦看到 S，攻方就假定面对强类型守方的修正概率为 $0.25/(0.25+0.75x)$。同样，观察到 S 后，认为守方为弱类型的修正概率是 $0.75x/(0.25+0.75x)$。

现在我们可以用攻方在观察到 S 后的概率修正信息来确定弱类型守方使得攻方在 I 和 NI 之间无差异的 x 值。如果攻方观察到 S，他选择 I 的期望支付为

$$\frac{0.25}{0.25+0.75x}(-10)+\frac{0.75x}{0.25+0.75x}5$$

攻方选择 NI 的期望支付为 0。若 $[0.25\times(-10)]+(0.75x\times5)=0$，或 $3.75x=2.5$，或 $x=2/3$，攻方就会发现选择 I 和 NI 的期望支付相等，于是愿意混合使用它们。若攻方看见 NS，他就肯定地推断守方为弱类型（只有弱类型守方才会选择 NS）并因而理所当然地选择 I。

迄今为止，我们已知道，弱类型守方以 2/3 的概率选 S 而以 1/3 的概率选 NS，就可以使攻方在 I 和 NI 之间无差异。然而，要使之成为均衡，则攻方观察到 S 后的混合策略概率也必须使弱类型守方在 S 和 NS 之间无差异。假设攻方看见 S 后选择 I 的概率为 y，则弱类型守方从 S 得到的期望支付为 $-9y-4(1-y)$；若他选择 NS，则肯定会带来攻方选择 I 的反应，所以弱类型守方将得到 -5。当 $-5y=-1$，或 $y=1/5$ 时，两种情况支付相等而弱类型守方愿意混合其策略。

最后，我们需要确认强类型守方肯定选 S。给定 $y=1/5$，选择 S 使其得到 $-14y-4(1-y)=-6$；选择 NS 则得到 -10。所以他宁愿选择规定的纯策略 S。

将所有的结论放在一起，它们的确满足了半分离均衡情况的所有条件。强类型守方始终选择 S；弱类型守方混合采取 S(在 2/3 的时间里）和 NS(在 1/3 的时间里）；而攻方观察到 S 后就运用贝叶斯法则修正信息，混合采取 I(在 1/5 的时间里）和 NI(在 4/5 的时间里），而观察到 NS 后则肯定采取 I。

9.5.4 一般理论

到目前为止，我们对攻—守博弈的分析仍局限于图 9—1 和图 9—2 所给出的

具体支付和概率值。考虑博弈更一般的形式是可以的，如图 9—3 就是用变量来表示支付的。这里，C 表示发送信号（S）的代价；F 表示战争的代价（若攻方侵略守方）；G 代表（攻方从侵略弱类型守方中）所得的收益。同样，p 表示守方为强类型的概率。对照看一下，先前的博弈中，F 是 10，G 是 5，而 p 是 0.25；C 开始是 6 然后是 4。

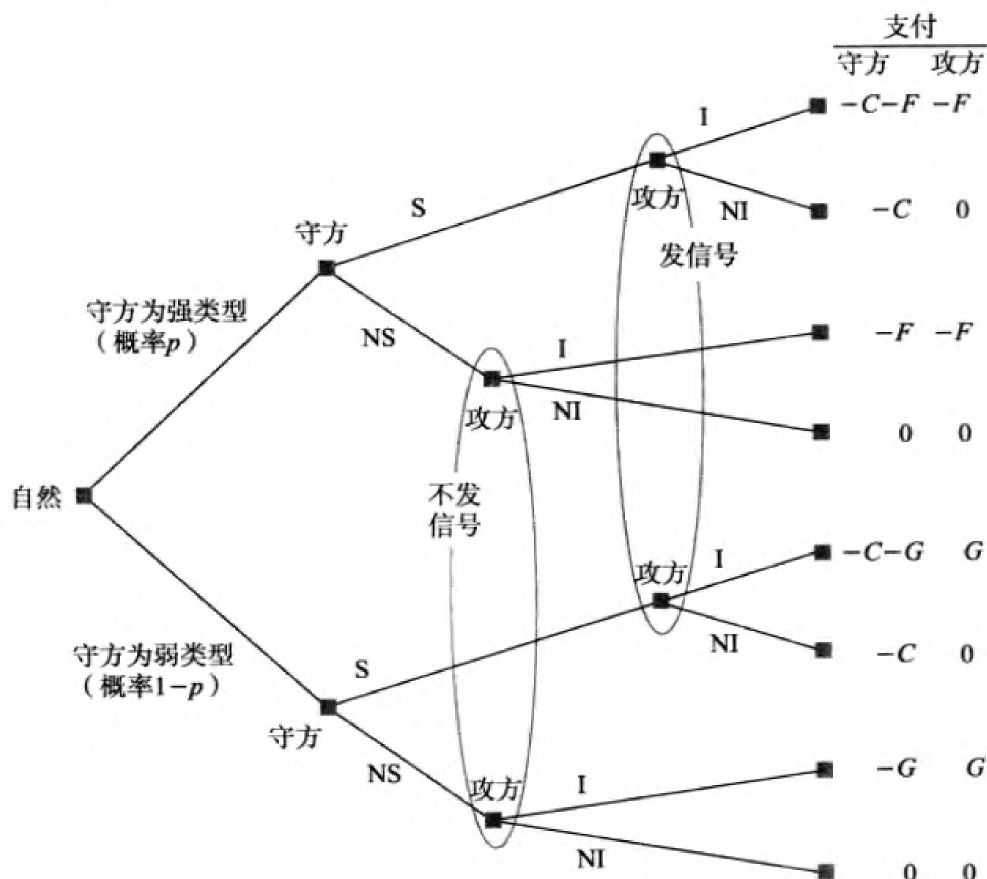


图 9—3 信号传递博弈（一般形式版本）

使用一般的支付结构，我们可以再进行一次 9.5.1 节至 9.5.3 小节的分析并推导参数 C，F，G 和 p 在不同均衡出现时的取值条件。不妨回想，我们在 9.5.1 小节中得到了一个分离均衡：强类型守方选择 S 而弱类型守方选择 NS，且攻方观察到 S 就选择 NI 而观察到 NS 就选择 I。按照同样的推理步骤，我们可发现，给定守方的策略，攻方的推断是正确的且其选择也是最优的，因为 $G>0$。强类型守方选择 S 导致对手选择 NI，从而获得 $-C$；选择 NS 导致 I 从而得到 $-F$。因此，只要 $-C>-F$，或 $C<F$ 时，守方选择 S 就是最优的。同样的计算表明，若 $-G>-C$，或 $G<C$ 时，弱类型守方选择 NS 就是最优的。把这些结果放在一起，我们发现图 9—3 的一般性博弈中，当 $G<C<F$ 时，分离均衡就会出现；注意，我们在图 9—1 中所选择的数字（$G=5$，$C=6$，$F=10$）是满足这个限制条件的。

针对其他各种可能的均衡类型，可以进行同样的计算。我们把结果列在表 9—6 中。该表列出了每种类型的均衡——分离均衡、混同均衡、半分离均

衡——满足 PBE 标准的策略集，以及使得那些策略集满足各类均衡的 PBE 标准的参数值。擅长代数的读者可以把它当做一个例子来计算。其他读者可以不去理会计算细节，而只需要留意表中列出的某些重要思想就可以了。

表 9—6　　信号传递博弈均衡的一般结果

均衡类型	均衡策略：守方 强类型	均衡策略：守方 弱类型	均衡策略：攻方	均衡存在条件
（完全）分离均衡	S	NS	若 S 则 NI，若 NS 则 I	$G<C<F$
	NS	S	若 S 则 I，若 NS 则 NI	永不存在
混同均衡	NS	NS	若 S 则不修正信息；若$(1-p)G>pF$ 则 I，若$(1-p)G<pF$ 则 NI	任何 G,C,F，或 p
			若 S 则假定为强类型，始终选择 I	$C>F,C>G$ 且$(1-p)G>pF$
	S	S	若 NS 则不修正信息；若$(1-p)G>pF$ 则 I，若 $(1-p)$ $G<pF$ 则 NI	永不存在
			若 NS 则假定为弱类型，若 I$(1-p)G>pF$ 则 I，若$(1-p)G<pF$ 则 NI	$C<F$，$C<G$ 且 $(1-p)G<pF$
半分离均衡	S	混合（以概率 x 选择 S）	用 p 和 x 修正信息；若 S 则混合策略（以概率 y 选择 I），若 NS 则 I	$C<G$ 且 $(1-p)G>pF$
	混合（以概率 x 选择 S）	NS	用 p 和 x 修正信息；若 NS 则混合策略（以概率 y 选择 I），若 S 则 NI	$C<F$ 且 $(1-p)G<pF$

首先，导致不同类型均衡的支付值和概率值的区间并非相互排斥的。因此，给定一个 C，F，G 和 p 的组合，有可能出现多重均衡。例如，若不正常地选择 S 被解读为两种类型守方均可能以相同概率犯下的错误，那么对于任何 C，F，G 和 p，任何类型守方都不发送信号而攻方总是选择侵略就是一个可能的混同均衡。所以，这样的均衡与那些要求特定的 C，F，G 和 p 值才能存在的均衡类型是可以共同存在的（coexist）。更为关键的是，同时满足 $G<C<F$（分离均衡要求）和 $C<F$ 且 $(1-p)G<pF$（半分离均衡要求）是可能的。$C<F$ 且 $(1-p)G>pF$ 也是可能的，并会同时得到混同均衡和半分离均衡。

我们已遇见过其他情形下的均衡多重性。例如，第 4 章和第 7 章的懦夫博弈有两个纯策略均衡和一个混合策略均衡。我们指出，为了从可能的均衡集合中选择一个结果，则其他的一些原理或程序——比如聚点、承诺机制等——是必要

的。当前情况也是如此。人们为此提出了诸多对完美贝叶斯均衡进行“精炼”(refinements)的方法，但那些方法在数学上都太复杂，并且对方法之选择也并非总是很明确。所以，我们把对均衡的精炼的研究留给更高级的博弈理论。

接下来讨论一下这个例子中完全分离均衡或半分离均衡下信号传递为什么有用。通常，不同类型守方之间总会有些差异，然而我们曾假设信号传递的直接成本 C 对于两类守方并无差异。其行为的差异产生于信念和行为的互动。例如，在完全分离均衡中，S 导致反应行为 NI 而 NS 导致反应行为 I。所以选择 S 和 NS 之间的支付差异，对于强类型守方是$(-C)-(-F)=F-C$，对于弱类型守方是$(-C)-(-G)=G-C$。由于这种情况下 $G<C<F$，所以 S 对于强类型守方是更好的选择但对于弱类型守方却是更糟的选择。

最后，根据我们的结果可以解决杰维斯悖论。杰维斯认为弱小的防守方更需要避免侵略，因此他发送强类型的信号可能被看做弱小的标志。这种情况的完全分离均衡表示在表 9—6 的表格第二行，但我们看到无论 C，F，G 和 p 为何值，该结果永远不会成为均衡。直观上看来，信号是代价沉重的，如果攻方按照与防守方相反的方式解读信号，那么理性的防守方会发现发送这样的信号不是最优的。

但是，半分离均衡中，如果守方为强类型则选择发信号，如果守方为弱类型则不发信号（表 9—6 中从表底倒数第二行）；这一均衡可以呈现出杰维斯论点的某些特征。弱类型守方会试图显露得比较强大但却只能部分地这样做。同样的故事发生在安德森（Elijah Anderson）所说的贫困街区的年轻人身上。他发现很多（但不是全部）衣着光鲜试图在主流社会取得成功的人仅仅只有伪装和模仿成功人士的能力，以便在“街头”（street）为他们提供一种安全感。[19]这是弱者部分地模仿强者的例子。

9.6 信号传递和甄别的一些例证

如我们在第 5 节所见，信号传递和甄别博弈的完美贝叶斯均衡解及其特征包含微妙的概念及计算。我们能否认为参与人可以正确地进行计算？这些博弈的均衡结果又有多少真实性？

有很多证据表明人们在概率计算尤其是关于新信息的条件概率计算方面非常糟糕。[20]基于观察到的行动修正其信息的过程中，这些计算是不可避免的。所以，我们有理由怀疑依赖于这种计算的均衡。虽然如此，经济学家从信号传递的实验结果得到的发现仍令人振奋。他们成功地发现了完美贝叶斯均衡的精炼结果，尽管获得这些精炼结果不仅需要观察均衡路径上的一系列行动从而及时修正信息，而且需要明确参与人如何从那些绝不会被率先采取的行动中获知信息。然而，实验的结论并非毫无争议，这很可能是实验设计的一些细节上的不同所致。[21]

虽然信号传递和甄别博弈的均衡看似复杂，但信号传递和甄别所起的作用却

非常简单，不过是用以套取信息而已。不同“类型”的参与人（在个人性格方面，或在参与的具体博弈方面，更一般地在博弈的支付方面拥有不同的信息），将会采取各自不同的最佳行动，所以他们的行动会真实地披露他们的类型。这一思想有诸多应用，以下是几个例子；而如果你沿着这些线索去思考，你会得到更多的例子。

（1）保险公司通常会提供一系列规定了不同自负额及共同保险的方案。自认为运气不好的人会偏好低自负及低共保的方案，而那些自认为运气不错的人则会偏好自己承担更多损失的方案。所以不同的风险等级有不同的最佳行动，保险公司让其顾客自行选择，以甄别和诱导出每位顾客的风险等级。

（2）风险资本家寻找发明家，投资他们的新产品，但是如何判断新产品的好坏？他们会衡量发明家在经济许可的情况下参与投资的意愿，所以承担损失的意愿变成了发明家对其产品获益是否有信心的可靠指标。

（3）轿车、电脑之类的耐用品，对于消费者而言很难评估其质量好坏，而每家厂商对自己的产品质量却很了解。如果产品质量是高的，则给予一个更长的保修期并不会花太大的代价。所以消费者就可以根据保修期长短来判断质量并做出购买决策。

（4）最后一个例子来自生物学。[22]在很多鸟类中，雄鸟通常有鲜艳而厚实的羽毛，以吸引雌鸟。人们可能认为雌鸟在寻找基因优良的雄鸟，这样它们的后代才能有优良的基因吸引异性。但是，为什么大而厚实的羽毛可以代表基因优良？人们也可以认为那是缺陷，因为太醒目的羽毛更容易被猎人发现，而且行动也不方便，很容易被抓获。这样的话为什么雌鸟要选择有缺陷的雄鸟呢？答案来自可置信信号传递的条件。尽管厚重的羽毛是个缺陷，但是只有强健且有速度的鸟才能承受，越弱的鸟越不能负担厚重的羽毛，所以厚重的羽毛的确传递了雄鸟体质的可靠信号。

正如你从这些例子或其他例子比如第 1 章的约会博弈小故事中可发现的，信息不对称随处可见，而其应对策略构成了博弈理论科学的一个重要部分。但是，策略的艺术也在日常生活中存在。我们希望这些思想可以激发你的兴趣去获取更多的相关知识，而不仅仅是读读我们这本入门教科书。基于这一目的，我们向大家推荐如下附加阅读书目。

社会学著作：Erving Goffman，*The Presentation of Self in Everyday Life*，revised edition（New York：Anchor Books，1959）。

生物学著作：Matt Ridley，*The Red Queen：Sex and the Evolution of Human Behavior*（New York：Penguin Books，1995）。

经济学著作：A. Michael Spence，*Market Signaling*（Cambridge：Harvard University Press，1974）。

还有来自经济学领域的两本研究生级别的教科书，那些雄心勃勃想了解更高水平的博弈理论的学子一定要读一读：Drew Fudenberg and Jean Tirole，*Game Theory*（Cambridge：MIT Press，1991），Chapter 6～8；and David Kreps，*A Course in Microeconomic Theory*（Princeton：Princeton University，1990），Chapter 13，16，and 17。

9.7 总　　结

具有私人信息的参与人可能想隐匿或披露其信息，而没有这些信息的人则试图诱取或回避这些信息。存在非对称信息时，行动胜于语言。要披露信息，则需要可置信的信号。信号传递只有在拥有不同信息的参与人发送（相同）信号具有不同代价时才有效。当直接询问不足以获取真实信息时，要获取信息，就需要甄别方案，这可能要求寻找到一个具体的行动。甄别只有在甄别机制诱导别人披露其真实类型时才有效；要分离不同类型的参与人，必须要有激励兼容条件。信息非对称的解决需要付出代价，有时候还会产生外部效应。

在不完美信息博弈均衡中，参与人不能只按照自己的信息采取最优行动，还必须根据对他人行动的观察进行推断（更新其信息）。这种类型的均衡就是完美贝叶斯均衡，它存在三种结果：混同均衡、分离均衡和局部分离均衡，这取决于支付结构以及对信息修正过程的规定。在某些参数区间，还可能存在多重类型的 PBE。

尽管完美贝叶斯均衡需要复杂的计算，但证据表明参与人还是有能力达到该均衡的。不同的实验设计会有不同的实验结果。生活中有大量的信号传递和甄别博弈例子，比如说保险公司提供的保险方案以及政府发行公债等。

关键术语

废话均衡（babbling equilibrium）
廉价磋商均衡（cheap talk equilibrium）
置信度（credibility）
激励相容约束（incentive compatibility constraint）
激励方案（incentive scheme）
道德风险（moral hazard）
局部信息披露均衡（partially revealing equilibrium）
参与条件/约束（participation condition/constraints）
完美贝叶斯均衡（perfect Bayesian equilibrium）
类型混同（pooling of types）
甄别机制（screening device）
甄别（screening）
自选择（self selection）
半分离均衡（semiseparating equilibrium）
类型分离（separation of types）
信号干扰（signal jamming）
信号（signal）
信号传递（signaling）

习　　题

1. 某知名品牌速溶咖啡电视广告中，一位男士在他的寓所招待一位女士。他想

请她来点咖啡及点心，当女士同意后，他走到厨房泡速溶咖啡同时假装高级（且昂贵的）咖啡机冲泡时的声音。正当他这样做的时候，另外一个房间传来女士的声音："我要看看咖啡机……"

运用非对称信息博弈的知识，评论这两个人的行动。请注意他们所使用的信号传递和甄别策略，并举出每个策略的具体情节。而在你看来，他们中谁是更厉害的策略高手？

2. "罗宾逊先生（Mr. Robinson）认为商学院是一种筛选机制——MBA 学位是雅皮士的最爱。但斯坦福商学院的最大特色也许就是，在上课之前就完成了筛选。只要能进去就一定能毕业，'他们不想让你被挂，他们要你成为最有钱的校友，回馈给学校很多钱'。但是有人怀疑：如果企业把选择年轻经理人的责任推给斯坦福招生办，那为什么不直接拆掉人事部门，换成斯坦福的招生小组，顺便去掉似是而非的 MBA 教育？浪费一大笔钱还有两年的青春，来证明自己能致力于商业管理，这值得吗？"（摘自迈克尔·刘易斯（Mikael Lewis）对皮特·罗宾逊（Peter Robinson）所著的《地狱惊幕：攻读 MBA》（*Snapshots from Hell: The Making of an MBA*）一书的评论，见《纽约时报》1994 年 5 月 8 日书评专栏。）基于你对不对称信息的认识，你将如何回答刘易斯的问题？
3. 向客户提供保险以保护其财产免遭盗窃或意外事故损失的企业，对其保险客户的行为深感兴趣。企业可使用某些激励方案防止并检测客户诈保或疏忽大意。请简要说明这些激励方案的基本思想。
4. 假设有两类电工：有能力的和无能力的。两类电工都可拿到资格证书，但是无能力者要拿到证书就必须付出额外的努力。有能力者需耗时 C 个月去通过资格考试，无能力者则需要两倍的时间。有资格证书的电工在合法承包商的工地工作每年可赚 10 万美元。没有资格证书的电工从事一些临时性工作每年只能赚到 2.5 万美元，也是合法承包商雇用他们。每类电工得到的支付为 $\sqrt{S}-M$，其中 S 是以千美元为单位衡量的薪水，M 是获取证书所耗的时间月数。请问 C 处于何区间可令有能力的电工选择证书信号，而无能力的电工不选择证书信号？
5. 在本章第 4 节经理努力程度问题的例子中，假设项目成功可为企业赚得 600 000 美元。现在其他条件不变（即经理只一般性努力时成功率是 60%，而要他付出额外努力就必须再给他 50 000 美元，等等），请证明：若基本薪酬不能为负，当项目成功的价值仅 400 000 美元时，企业支付低工资、要求经理做一般努力是更好的选择。
6. 某个经济中有两类工作：好工作和差工作；有两类工人：有资格证书者和无资格证书者。经济中 60%的工人持有证书，40%的工人无资格证书。每类工人从事差工作都可产出 10 单位；而有资格的工人从事好工作可产出 100 单位，无资格的工人从事好工作产出为 0 单位。假设公司对各类工人有足够多的需求而必须支付与其产值相等的工资。

公司在雇用工人之前，无法直接观察到工人的类型，但是有资格证书的工人以资格证书作为信号传递给公司。有资格证书的工人接受 n 级教育的代

价是$n^2/2$；而无资格证书的工人，该代价是n^2。这些代价跟产出采取同样的度量单位，且n必须是整数。

(a) 达到分离均衡的最小的n值是多少？

(b) 现在假设不能进行信号传递，则两类工人会从事何种工作？薪水各为多少？谁会因此而获益而谁会因此而受损？

7. Mictel公司在个人电脑生产上垄断市场，该公司可生产两类电脑：低端机与高端机。购买者人数是P，他们分作两类，分别是临时用户和经常用户。临时用户占总人数比例为c，经常用户比例为$(1-c)$。

两类电脑的生产成本以及两类不同用户的收益，由下表给出，单位是千美元。

		成本	用户的收益	
			临时用户	经常用户
电脑类型	低端机	1	4	5
	高端机	3	5	8

每类用户计算他使用不同电脑的净支付（收益减去价格），并且购买有最高净支付的电脑，前提是净支付不能为负。如果两类电脑带给用户的净支付相等，则用户会选择高端机；如果两类电脑带给用户的净支付为负，用户就不购买。而Mictel想要使期望支付最大化。

(a) 如果Mictel无所不晓，则它会知道前来购买者的类型，此时公司就能提供一种电脑给他并以不二价卖给他。请问面对不同类型的购买者，它会以什么价格提供什么电脑？

事实上，Mictel并不知道顾客的类型，它只是提供目录让购买者选择罢了。

(b) 首先，假设公司只生产低端机，并以价格x销售。请问使利益最大化的x值是什么？答案跟c有什么关系？c是临时用户的比例。

(c) 其次，假设公司只生产高端机，并以价格y销售。请问使利益最大化的y值是什么？答案跟c有什么关系？c是临时用户的比例。

(d) 最后，假设公司生产两种类型的电脑，以x价格卖出低端机，以y的价格卖出高端机。如果公司要让临时用户购买低端机，让经常用户购买高端机，则什么样的x和y值才满足激励相容约束？此时公司的期望利益是多少？让公司期望利益最大化的x值和y值是多少？答案跟c有什么关系？

(e) 把以上答案加以比较，Mictel公司要采取怎样的生产及价格策略？答案跟c有什么关系？

8. 费利克斯（Felix）和奥斯卡（Oscar）两人玩一种简单的扑克牌。每人开始有8美元作为赌本。然后大家分别抽一张牌，大或小的概率各占1/2。每人看到自己的牌但是不知道对方的牌。

然后费利克斯决定开牌或者加注（额外加赌4美元）。若他选择开牌，则两张牌被翻开并加以比较。如果两张牌不一样，则大者胜，并赢得全部赌金

16 美元（其中包括他自己的 8 美元，所以他实际支付为 8 美元）；输家亏损 8 美元。若两张牌一样大，则各自拿回 8 美元（支付为 0）。

如果费利克斯选择加注，那么奥斯卡决定关牌（放弃）或者跟牌（亦跟注 4 美元）。若奥斯卡关牌，则费利克斯获胜（无论他的牌大小如何）并拿走全部赌金；若奥斯卡选择跟牌，则两张牌翻开比较。这一过程与上一段落所说的一样，只不过现在的赌金变得更多了。

(a) 以扩展式表示这个博弈（请注意信息集）。

如果将该博弈写成标准式，费利克斯有四个策略：(1) 总是开牌（简写为 PP），(2) 总是加注（RR），(3) 在自己的牌大时加注而牌小时开牌（RP），(4) 牌大时开牌而牌小时加注（PR）。同样，奥斯卡也有四个策略：(1) 总是关牌（FF），(2) 总是跟牌（SS），(3) 在自己牌大时跟牌而牌小时关牌（SF），(4) 牌大时关牌而牌小时跟牌（FS）。

(b) 证明费利克斯的支付如下表所示：

		奥斯卡			
		FF	SS	SF	FS
费利克斯	PP	0	0	0	0
	RR	8	0	1	7
	RP	2	1	0	3
	PR	6	−1	1	4

（对于以上 16 个单元格，你必须计算四种不同的抽牌组合，然后求期望值。）

(c) 尽量剔除劣策略，在剩下的表中找出混合策略均衡，以及费利克斯在均衡中的预期支付。

(d) 运用信号传递和甄别理论知识，解释为什么均衡时有混合策略？

9. [讨论] 健康保护体制设计在几个方面涉及了信息和策略的问题。用户——潜在的或事实上的病人——对自己的健康状态、生活方式等比保险公司有更多的信息。而健康保护服务提供方——医生、医院等——也比病人和保险公司更清楚病人需要什么。医生对自己的医术和责任心了解更多，医院对设施了解更多。虽然保险公司能从过去的病例中得到手术及治疗的统计结果，但这些统计结果会受到许多不可观察或随机因素的影响，所以观察医院的设施和医生的医术并不能完全推断结果。而制药公司也比其他人更清楚药的效力，一般而言，它们没有自然而然的动机与人分享这些信息。所以健康保护体制设计必须面对这些问题，找出最优的可行方案。

请从策略的角度讨论不同支付方案的优点：医生采取服务计费或人头计费，病人缴年费或者按次计费等方式，哪一种对医疗需求者最有利？哪一种对医疗提供者最有利？也请思考一下私人投保以及从国家税收支付保费的相对优点。

10. [选做] 某老师想要知道学生对他们自己的能力有没有自信心。他提出以下方案："当你们回答问题后，请说出你估计答对的概率。然后我会跟你的答案对比，假设你给的概率是 x，如果你答对了，则成绩就是 $\log x$；如果答错，则成绩是 $\log(1-x)$。"请证明，这个方案可以诱导出学生对自己的真正估计；并且，若真正概率为 p，请证明学生估计的概率 $x=p$。

【注释】

[1] 非对称信息经济学的理论先驱 George Akerlof，Michael Spence 和 Joseph Stigliz 因为这些贡献获得了 2001 年诺贝尔经济学奖。

[2] 提醒一句：不要混淆甄别（screening）和信号干扰（signal jamming）。用通俗的话说，"screening"一词可以有不同的意思。博弈论中使用这个词的意思是检验（testing）和审查（scrutinizing）。所以信息劣势方使用甄别机制去发现信息优势方的信息。"screening"的另一个意思即隐匿遮蔽（concealing）——博弈论术语是信号干扰（signal jamming）。所以信息优势方使用信号干扰行动去阻止信息劣势方正确地从行动中推断出真相。

[3] 当一个参与人有三个或更多的行动时，讨论就会稍微麻烦一点。若埃弗特选择行动 DL，CC 和 Lob，并说"我选 DL"，则纳芙拉蒂洛娃应当不相信并且也不能解读为 DL 完全被排除。根据这一逻辑又会回到同样的结论，即纳芙拉蒂洛娃应当对埃弗特的话毫不理会。

[4] 第二阶段中，双方都正好选择与哈里消息相反的地点之可能性如何？这也是第二阶段子博弈的一个纳什均衡，所以有可能存在"不当的"廉价磋商均衡。这种情况下，哈里在第一阶段的最优行动是说"我去本地咖啡店"，这样他仍可获得其偏好的结果。

[5] 发展了局部沟通理论的研讨论文来自 Vincent Crawford and Joel Sobel，"Strategic Information Transmission，" *Econometrica*，vol. 50，no. 6，November 1982，pp. 1431-1452。初步的研究和深入研究工作的综述参见 Joseph Farrell and Matthew Rabin，"Cheap Talk，" *Journal of Economic Perspectives*，vol. 10，no. 3，summer 1996，pp. 103-118。

[6] 我们只要求选择某类型候选者的支付至少要与选择其他类型候选者的支付一样高，而不是严格地高。但是，只要人们愿意，严格大于的结果可以逼近等于的结果，所以这一假设并无实质性的缺陷。

[7] 在其他背景下，相应的选择不一定要求必须是整数——例如，金钱或时间的总和，由此满足这些条件的数值可以是一个连续区间。

[8] 正如我们将在 12.5 节看到的，C 类对 A 类强加了一种负外部效应。

[9] 注意，类型混同跟我们在本章附录中提到的风险混同（pooling of risks）不一样。后者发生于团体中每个人都面临不确定性的时候。由于每个人都面临不确定性，于是大家同意将资产联合起来并分享资产回报。只要他们的风险不是完全正相关的，这样的安排就可以降低每个人的风险。而类型混同中，参与人的本质（例如能力）不同，但仍接受相同的结果或待遇（比如工作和薪水）。注意，类型混同正好跟类型分离相反，后者指不同本质的参与人会得到不同的结果，所以结果可以完全揭示参与人的本质。

[10] 这个概率可这样看待：从潜在的博弈参与人中随机抽取一些守方样本，其中强类型占守方样本的 1/4。

[11] 对此进行分析的文献包括 Robert Jervis，*The Logic of Images in International Relations*，Princeton：Princeton University Press，1970；Barry Nalebuff，"Rational De-

terrence in an Imperfect World," *World Politics*, vol. 43, no. 2, April 1991, pp. 313 - 335; James Fearon, "Signaling Versus the Balance of Power and Interests," *Journal of Conflict Resolution*, vol. 38, no. 2, June 1994, pp. 236 - 269; 以及 Robert Powell, "Uncertainty, Shifting Power, and Appeasement," *American Political Science Review*, vol. 90, no. 4, December 1996, pp. 749 - 764。

[12] 这一方面研究的综述参见 Robert Wilson, "Strategic Models of Entry Deterrence," *Handbook of Game Theory*, vol. 1, ed. Robert Aumann and Sergiu Hart, Amsterdam: North-Holland, 1994, pp. 309 - 325; 以及 Kyle Bagwell and Asher Wolinsky, "Game Theory and Industrial Organization," *Handbook of Game Theory*, vol. 3, ed. Robert Aumann and Sergiu Hart, Amsterdam: North-Holland, 2002, pp. 1851 - 1895。

[13] 例如，非洲热带草原上小羚羊看见猎豹逼近的时候，它会在适当位置不断跳起在 6 英尺以上。其目的是给猎豹传递信号，小羚羊太强壮了，要追赶它只会无功而返。见 Fernando Vega-Rodando and Oren Hasson, "A Game-Theoretic Model of Predator-Prey Signaling," *Journal of Theoretical Biology*, vol. 162, no. 3, June 1993, pp. 309 - 319。

[14] Elijah Anderson, *Code of the Street*, New York: Norton, 1999, pp. 23, 73。

[15] 见本章附录对贝叶斯法则的解释。

[16] 人们也可能会认为，若某类型的守方总是伪装成另一种类型，则攻方就可以在观察到行动 NS 时推断守方为强类型，观察到 S 就推断守方为弱类型，此时分离结果也是可能的。然而，这样的一套策略不是我们所规定的博弈均衡。我们在 9.5.4 小节中考察这种可能性。

[17] 攻方完备的行动计划必须详细规定他对守方所有可能选择的反应，即使是那些在均衡中不大会发生的选择也要纳入行动计划。观察到本来从不应发生（即概率为 0）的事件之后，并不能用贝叶斯法则来进行概率修正。这里我们要做的是说明攻方可以采取的两个合理的惯常假设。大量的文献讨论了不同假设的合理性和结果，但是我们把这些留给更高级的分析。要进一步了解这一主题，可参考 Drew Fudenberg and Jean Tirole, *Game Theory*, Cambridge: MIT Press, 1992。

[18] 这里又会碰到配置混合策略均衡（包括两类守方的混合策略概率）的其他可能方式。但我们关注存在半分离均衡的情况，而把其他的可能均衡及其存在条件的推导留给意欲对此问题进行深入分析的人们。

[19] 参见 Anderson, *Code of the Street*, p. 36。

[20] Deborach J. Bennett, *Randomness*, Cambridge: Harvard University Press, 1998, pp. 2 - 3, and Chapter 10。见 Paul Hoffman, *The Man Loved Only Numbers*, New York: Hyperion, 1998, pp. 233 - 240, 本书讲到了几个概率论专家，包括才华横溢著作等身的数学家 Paul Erdös，他们算错了一个很简单的概率题，跟他们解释后竟然还不知道自己的错误。

[21] Douglas D. Davis and Charles A. Holt, *Experimental Economics*, Princeton: Princeton University Press, 1995。对这些实验的评论和讨论参见该书第 7 章。

[22] 参见 Matt Ridley, *The Red Queen: Sex and the Evolution of Human Nature*, New York: Penguin Books, 1995, p. 148。

附录　信息与风险

当参与人在博弈中拥有不同数量的信息时，他们就会使用一些机制去探测对手的私人信息。我们在 9.3 节看到，某些时候可以依靠直接沟通产生廉价磋商均衡。但更多的时候，参与人需要通过观察对手的行动来确定对手的信息。所以他们必须通过对手的行动或已观察到的结果来估计潜在信息的概率。这种估计要用相对复杂的概率法则来运算，我们将在本附录第 1 节详细讨论概率计算过程。

同样，当某行动结果具有风险性时，则新的策略问题，即必须对付风险操纵本身，就会产生。一个规避风险的参与人可以通过设计降低风险的策略来得到好处，尤其是当两个或更多参与人可以从联合或组合其风险中受益的时候。在某些博弈中，尤其是参与人相互竞争尔虞我诈的时候，其针对风险的策略互动极为有趣。我们将在本附录第 2 节讨论这个问题。

9A.1　从结果观察中进行概率推断

7A.1 节给出的事件概率运算规律，尤其是结合律，在参与人具有不同信息时计算支付非常有用。在非对称信息博弈中，参与人试图通过观察对手的行动来发现其类型，所以他们必须根据观察到的行动和结果去推断其潜在信息的可能性

(likelihood) ——估计其概率。

理解上述问题的最佳方式就是举例子。假设1%的人有遗传缺陷可能导致某种疾病。某项检验能以99%的准确性发现这种遗传缺陷：如果有缺陷，该检验有1%的时候不能发现，并且在没有缺陷的时候该检验也有1%的时候错误地报告有缺陷。我们感兴趣的是检验报告显示有缺陷时该人真有缺陷的概率。也就是说，我们无法直接观察一个人的遗传缺陷（潜在条件），但是我们可以观察缺陷检验结果（后果）——除非该检验不是缺陷的完美指标；给定我们的观察，我们能在多大程度上确定潜在条件是现实存在的？

我们可以用简单的数字运算来回答这个问题。假设人口总数为10 000人，其中100（1%）人有缺陷而9 900人没有。假设所有人都去做这个检验，则100个有缺陷的人检验结果显示有缺陷的人数为99；而9 900个无缺陷的人检验结果（错误地）显示有缺陷的人数为99。总共有198人检验显示有缺陷，其中一半人的检验结果是对的，一半人的检验结果是错的。如果随机地抽取一个检验结果显示有缺陷的人，由于检验结果正确和结果错误的可能性是一样的，所以报告显示有缺陷的人事实上存在缺陷的风险仅为50%（这是为什么稀有条件检验必须良好设计以降低产生“错误报告”的失误率的原因）。

对此类问题的一般情况，我们用所谓的贝叶斯定理（Bayes' theorem）代数公式来构造问题并进行计算。为了这样做，可将我们的例子进行一般化，允许两种可选的潜在状态 A 和 B（例如有缺陷或无缺陷），以及两个可观察结果 X 和 Y（例如检验结果为有缺陷或无缺陷）。假设，在缺乏有关人口整体的任何信息的情况下，A 存在的概率为 p，因此 B 存在的概率为 $(1-p)$。当 A 存在时，则观察到 X 出现的概率为 a，因此观察到 Y 出现的概率为 $(1-a)$［用第7章附录的语言来说，a 是以 A 为条件的 X 的概率，$(1-a)$ 是以 A 为条件的 Y 的概率］。同样，当 B 存在时，观察到 X 出现的概率为 b，因此观察到 Y 的概率为 $(1-b)$。

这样的描述告诉我们，可能有四种事件组合会出现：(1) 存在 A 并观察到 X；(2) 存在 A 并观察到 Y；(3) 存在 B 并观察到 X；(4) 存在 B 并观察到 Y。运用修正乘法律，我们可得到四种组合的概率分别为：pa，$p(1-a)$，$(1-p)b$ 和 $(1-p)(1-b)$。

现在假设 X 已被观察到，即一个人进行了遗传缺陷检验并且结果显示有缺陷。然后我们把注意力集中在前面四个概率的子集上，即第一个和第三个，两者都涵盖观察到 X 的情况。两个概率加总为 $pa+(1-p)b$，这正是 X 被观察到的概率。在 X 被观察到的这个概率子集里，A 也存在的概率刚好是 pa。所以我们现在知道了只观察到 X 的可能性，以及观察到 X 且又有 A 存在的可能性。

但我们更感兴趣的是，给定我们已观察到 X，那么 A 存在的可能性有多大——也就是说，若某个人的检验报告显示有缺陷，那这个人真正有遗传缺陷的概率是多少？这个计算是最绕的一个。使用修正的乘积律，我们知道 A 和 X 皆发生的概率，等于 X 单独发生的概率乘以以 X 为条件 A 发生的概率，即我们在下面给出的最后概率。使用先前算出的“A 和 B”发生的概率以及“只有 X”发生的概率，我们得到

$$\text{Prob}(AX)=\text{Prob}(X)\times\text{Prob}(A|X)$$
$$pa=[pa+(1-p)b]\times\text{Prob}(A|X)$$
$$\text{Prob}(A|X)=\frac{pa}{pa+(1-p)b}$$

这个公式向我们给出了给定观察到 X 后 A 发生的概率评估。这个结果就是贝叶斯定理（或法则或公式）。

在检验遗传缺陷的例子中，我们有 $\text{Prob}(A)=0.01$，$\text{Prob}(X|A)=a=0.99$，$\text{Prob}(Y|A)=b=0.01$。我们可将此三值代入贝叶斯公式，得到

给定检验报告显示有缺陷而实际上也存在缺陷的概率 $=\text{Prob}(A \mid X)$

$$=\frac{(0.01)(0.99)}{(0.01)(0.99)+(1-0.01)(0.01)}$$
$$=\frac{0.0099}{0.0099+0.0099}$$
$$=0.5$$

先前，我们是在列举所有可能情况的基础上来计算条件概率的，概率代数学则使用贝叶斯法则来确定条件概率，其好处是，一旦有了这个公式我们就可以机械地应用它，而且不必一一列举所有可能性并确定每种可能的概率。

我们将贝叶斯法则表示在表 9A—1 中，它可能比前面的公式更容易被记住。表的行表示可能存在的真实状态，例如“有遗传缺陷”或“无遗传缺陷”，在这里我们也只有两种可能 A 和 B，但是可以方便地进行推广。表中的列表示观察到的事件——例如“检验有缺陷”和“检验无缺陷”。

表 9A—1　　贝叶斯法则

		观察结果		行的合计
		X	Y	
真实状态	A	pa	$p(1-a)$	p
	B	$(1-p)b$	$(1-p)(1-b)$	$1-p$
列合计		$pa+(1-p)b$	$p(1-a)+(1-p)(1-b)$	

表中的每个单元格表示真实条件和观察结果的全部组合概率；它们也正好是先前四个可能组合的概率。右边最后一列表示的是前两列在对应行中的合计，该合计就是每一种真实状态的总概率（例如，A 的概率是 p）。最后一行表明前两行对应的列的合计，该合计给出了每一个观察结果所发生的概率。例如，X 列最后一行的数字就是观察到 X 的全部概率，包括 A 为真实状态（真正有缺陷）和 B 为真实状态（没有缺陷）的情况。

给定一个特定的观察结果，要寻找一个特定状态的概率，那么根据贝叶斯法则，我们应将该组合的真实状态和观察结果所对应的单元格中的数字除以该观察结果所在列最后一行的合计，比如，$\text{Prob}(B|X)=(1-p)b/[pa+(1-p)b]$。

练习：在遗传检验例子中，假设检验显示无缺陷（观察到 Y），那么一个人

真的没有遗传缺陷（B 存在）的概率是多少？运用贝叶斯法则计算此概率，然后用列举计算法检查你的答案是否正确。

9A. 2 控制和操纵风险

9A. 2. 1 降低风险的策略

假想你是一个靠天吃饭的农夫，风调雨顺的年成你可收入 15 000 美元，天气糟糕的年成你只能收入 5 000 美元。两种天气的可能性是相等的（即各自概率为 1/2，或 50%）。所以你的预期收入是 10 000 美元，但是这个平均收入有相当大的风险。

如何才能降低你面临的风险？你可能会尝试种植一些不那么受天气影响的作物，但假设你该做的都做了，现在你只有将自己的收入风险与他人的收入风险结合起来才能降低风险。有两种办法可达到这一目的：混同风险（pooling）或交易风险（trading）。

我们从**混同风险**（pooling risks）开始。假设你的邻居面临同样的靠天吃饭的风险，但是天气好坏的结果跟你正好相反（假设你们分别住在岛的两端，而风雨只会影响某一端）。学术圈内用专有名词相关性（correlation）来衡量两个不确定数量之间的一致性——在当前的讨论中，就是两个农夫的风险之间的一致性。我们说你的风险与你邻居的风险是**负相关**（negatively correlated）的。不管天气如何，你们共同得到 20 000 美元，也就是完全无风险的。你们可签立合约确保各自得到 10 000 美元：当你收成好时，答应给他 5 000 美元；当他收成好时，他答应给你 5 000 美元。因为你们对风险进行了混同，所以就消除了风险。

即使没有这么完全的相关性，混同风险也是有好处的。假设农夫的风险是独立的，就如同掷两个不同的硬币来决定两个人的天气好坏，那么就有四种可能结果，每种结果出现的概率是 1/4。第一种结果，双方都遇到好天气，所以两人总共得到 30 000 美元；第二种和第三种结果是，一方遭遇好天气，另一方遭遇坏天气，所以总共得到20 000美元；第四种结果是，双方都遭遇坏天气，所以总共只得到10 000美元。面对四种可能结果，如果你们签立合约分享总收入，则每个人有 1/4 的概率得到 15 000 美元，1/4 的概率得到 5 000 美元，1/2的概率得到 10 000美元。如果不签合约，每个人分别有 1/2 的概率得到 15 000 及 5 000 美元。所以合约降低了两个极端出现的概率，从1/2降到 1/4，且增加了中间结果的概率，从 0 增加到 1/2。换句话说，这份合约降低了你们每个人的风险。

事实上，只要两个农夫的收入不是完全**正相关**（positively correlated）的——也就是双方的天气好坏都一样，使得其收入也一样——则他们就可通过混同（pooling）来降低风险。而且，如果有两个以上的农夫面临某种程度的独立风险，则更可以通过签立合约大幅度降低每个人的风险。

降低风险，正是保险的基础：把许多类似但是相互独立的风险混合在一起，

当你遭受大量损失时保险公司来做赔偿。降低风险也是分散投资的基础：把你的财产分开来投资，让它们分别有不同程度的风险，则可以降低总风险。

然而，通过混同风险降低风险有其局限性。最简单的情况是，混同的承诺未必可信——收成好的农夫不一定想将他的收入拿出来分享。不过这很容易克服，只要有司法系统就可以强制履行合约，因为法官会观察农夫的收成，决定给谁多少钱。人们付给保险公司的保费，就是一种预先承诺，就好像把钱交给第三方托管，而保险公司为了维持声望，一定会按照约定给付赔偿金。然而，如果结果不能被他人察觉，就会有人假装受到损失而向保险公司诈保，或是某人的风险受到其本身的行为影响，则他在知道被保险后可能变得粗心大意，这被称作**道德风险**（moral hazard）。或者，如果保险公司不知道其风险大小（比如健康险），则它将面临风险越大的人投保越多，使得每个人的平均保费提高，这称作**逆向选择**（adverse selection），我们很快会再详细讨论。在这里，我们只是指出如何处理这些问题，保险公司通常只提供部分补偿，其他部分的损失要你自己负担，你负担的部分由自负额及共同保险组成。

接下来考虑**交易风险**（trading risk）。假设你是农夫，还是面对同样的天气风险，但现在你的邻居有稳定的收入 10 000 美元，你面临很大的风险但邻居却没有，他可能愿意承担你的小部分风险，如果风险的交易价格对于你们二人都还可以接受的话。

为了理解风险价格的计算，需要引入**风险规避**（risk aversion）的概念。假设你们二人都是风险规避的，在第 5 章附录中我们看到，你对风险的态度可用凹函数（a concave scale）将金钱收入转换为“效用”数量。在该附录中，我们曾用平方根作为效用函数的简单例子，这里继续沿用。假设你预先支付给邻居 100 美元，交换条件是：如果你收成不好，他给你 2 000 美元；但如果你收成好，则你给他 2 000 美元。换句话说，他用 2 000 美元的风险，跟你交换 100 美元的预付费。如果你收成好，他的最后收入会变成 10 000＋100＋2 000＝12 100 美元；如果你收成不好，他的收入变成 10 000＋100－2 000＝8 100 美元。所以他的期望效用为

$$0.5\times\sqrt{12\ 100}+0.5\times\sqrt{8\ 100}=0.5\times110+0.5\times90=100$$

这个效用等于他不跟你交易时的效用 $\sqrt{10\ 000}=100$，所以有没有交易对于他而言是一样的。而你会不会预付 100 美元跟他交易 2 000 美元的风险呢？

没有交易时你的风险收入效用为

$$\begin{aligned}0.5\times\sqrt{5\ 000}+0.5\times\sqrt{15\ 000}&=0.5\times70.7+0.5\times122.5\\&=96.6\end{aligned}$$

如果你付给邻居 100 美元，当你的收成差时他给你 2 000 美元（所以他剩下 8 100 美元），所以你会得到 6 900 美元。反过来，当你收成好时，你要给他 2 100 美元（预付额 100 美元加上你从中拿出的 2 000 美元），所以他得到 12 100 美元以补偿风险，你剩下 12 900 美元。所以你的期望效用是

$$\begin{aligned}0.5\times\sqrt{6\ 900}+0.5\times\sqrt{12\ 900}&=0.5\times83.0+0.5\times113.6\\&=98.3\end{aligned}$$

所以，交易让你的效用更高，但对于你的邻居来说，却没什么两样。

你的邻居也许会建议另一种交易，让他有更高的效用，但让你的期望效用跟原来的 96.6 一样，但也有其他交易会让你们同时获得更高的期望效用。所以交易的范围很广，有一些对你有利，有一些对他有利，你们可以谈判，结果取决于双方的谈判能力。我们将在第 17 章细述谈判理论。如果有很多人在交易，则非个人的市场（将在第 18 章加以讨论）将决定风险的价格。这些交易的共同之处是，借由合理的价格，使风险低者降低风险高者的负担，以达到互利的效果。

事实上，风险价格及风险市场的概念，正是现代经济中几乎所有的金融安排之基础。股票、债券以及其他诸如衍生品之类的复杂金融工具，都是用来将风险分散给那些愿意以最低要价购买的人的。许多人认为金融市场纯粹是冒险的一种形式。从某方面来看的确如此，但风险越小的人越冒险，也许因为他们把风险分散了，所以他们可以尝试更冒险的企业计划，而金融市场也因风险交易而促进了企业的发展。

当然，这些市场也受到道德风险、逆向选择甚至直接欺骗的限制，所以不能完全做到分散风险。例如，公司的经理通常必须承担他们决策的部分风险，由于持有公司股票或由于其他机制，他们有努力为公司赚钱的动机。

9A. 2. 2 利用风险

风险不一定是坏的，在某些场合它甚至可以成为你的优势。不过，讨论一般理论会让我们离题太远，所以我们只提供一个利用风险的简单例子，也就是**期权**(option)。

期权给予股票持有者以买或卖的行动权利，但他不受必须买或卖的约束。例如，股票持有者可以在预设价格下买到股票，比如 110，而当前的股票价格是 100，而且有同样的概率上升到 150 或下降到 50。如果是前者，股票持有者会行使其权利，得到利益 40；如果是后者，他可以放弃权利，得到 0，所以他的期望利益是

$$0.5\times40+0.5\times0=20$$

如果不确定性更大，即股票有同样的概率上升到 200 或下降到 0，则该期权的期望利益是

$$0.5\times90+0.5\times0=45$$

所以股票潜在的不确定性越大，期权的价值就越高。当然，市场也能意识到这一点，所以购买这种期权的价格本身也随市场的波动而改变。

9A. 2. 3 在竞赛中操纵风险

在 9A. 2. 1 小节，农夫面临的风险是天气影响，它而不是其他农夫造成的。

而相反的情况是，博弈参与人可以影响自己和他人所面临的风险，从而他们可以策略性地操纵这种风险。一个基本的例子是公司间开发新信息或新产品的竞争，还有许多体育比赛也具有相同的特征。

体育比赛和相关的竞赛，其结果由技能和运气一起决定。你取胜的条件是

你的技能＋你的运气＞对手的技能＋对手的运气

或

你的运气－对手的运气＞对手的技能－你的技能

上述不等式左边是你的"运气剩余"(luck surplus)，简称 L，它是不确定的；假设其概率服从正态分布，或如图 9A—1 的钟形曲线。在横轴上的每一点，其对应到曲线上的高度代表着该 L 值的概率[①]，所以曲线在两点间覆盖的面积等于 L 在该两点间的概率。假设你的对手技能更好，所以你处于劣势。你的"技能赤字"(skill deficit)——也就是对手的技能减去你的技能——因而是正的，在图上以 S 点表示。如果你的运气剩余 L 超过技能赤字 S，则你将获胜，所以 S 点右边的面积（以阴影表示）代表你获胜的概率。如果你让钟形曲线更平坦，则 S 点右边的面积将增加。因此若你处于劣势，你应该采取曲线较平坦的策略。相反，如果你有优势，就应该设法减少博弈的不确定性。

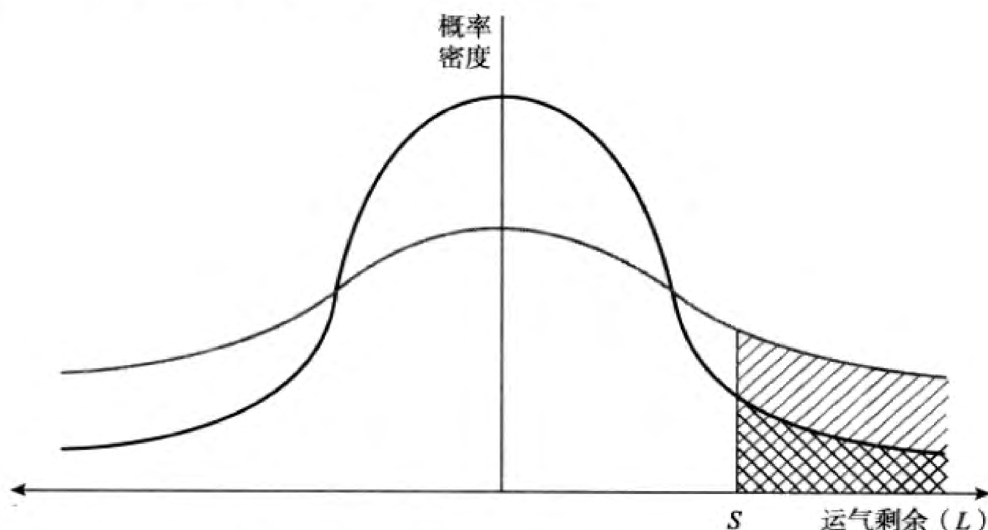


图 9A—1　更大的风险对取胜机会的影响

所以我们会看到比赛中处于劣势的人会使用冒险策略，因为"这是他们唯一反败为胜的机会"，而处于优势的人会采取安全策略。基于这个原则我们对你有个实践建议：如果你想要挑战网球打得好的人，那就选个风大的日子吧。

你不仅可以在策略中操纵风险，更可以操纵对手风险的相关性。领先的选手会尽量选择更高的正相关，这样一来，不论他的运气好坏，对手的运气也会跟他一样，这样他就可以继续保持领先。相反，落后的选手会选择负相关。有个著名

① 原文如此，疑应为"代表着该 L 值的概率密度"。——译者注

的例子，在两人帆船竞赛中，落后者应选择跟领先者不同的航道，而领先者应该顺着落后者的航道。[1]

练习：在前面的风险交易的例子中，你有一笔风险收入，好天气（概率0.5）可收入15 000美元，差天气（概率0.5）可收入5 000美元；你的邻居有稳定收入10 000美元。我们推导出了保持对方无差异而你效用提高的交易方案。请构造一个类似方案，该方案中你的邻居让你保持在最初的预期效用水平，而他的效用比不参与交易时更高。

9A.3 总　结

如果参与人在博弈中信息不对称，他们可以尝试根据对其行动或结果的观察来推断隐蔽的潜在条件概率。贝叶斯定理提供了进行这种推断的公式。

当博弈参与人对风险持不同的态度或拥有不同数量的信息时，操纵风险和信息的策略行为就大有用武之地。参与人可通过交易或混同（组合）来降低其风险，尽管后者因道德风险和逆向选择而变得更复杂。风险在某些竞赛中是有价值的，参与人可通过操纵它而获得好处，当然这要取决于博弈的环境。

关键术语

逆向选择（adverse selection）
混同风险（pooling risk）
贝叶斯定理（Bayes' theorem）
正相关（positively correlated）
负相关（negatively correlated）
交易风险（trading risk）
期权（option）

【注释】

[1] 这个例子可参考 Avinash Dixit and Barry Nalebuff, *Thinking Strategically*, New York: Norton, 1991, pp. 10-11。若希望了解更多的一般理论分析，请参考 Luis Cabral, "Football, Sailing, Tennis, and R&D: Dynamic Competition with Strategic Choice of Variance and Covariance," *Journal of Economics and Management Strategy*, Vol. 12, no. 1, Spring 2003, pp. 139-150。

第 10 章　策略性行动

一场博弈可由参与人的选择及可行的行动具体说明，参与人行动的次序及他们的支付都取决于他们的决策。在第 6 章，我们可以看到从序贯行动变为同时行动或者从同时行动变为序贯行动是如何改变博弈的结果的。增加或者减少参与人的行动，或者改变某些结点上或支付表中某些组合的支付，也会改变结果。除非博弈的规则是由权威制定的，否则每个参与人都有操纵规则的动机，以产生对他们自己更有利的结果。我们把这种操纵博弈的方法称为**策略性行动**（strategic moves），这就是本章的主题。

策略性行动会改变原来博弈的规则，从而产生一个新的两阶段博弈。从这个意义上说，策略性行动与 9.2 节中说到的直接信息沟通有相似的地方。经过策略性行动后，第二阶段虽然仍是原来的博弈，但在行动次序和支付上有些改变，而在存在直接沟通的博弈中却没有这种改变。存在策略性行动的博弈中，第一阶段会指明在第二阶段要做些什么。不同的第一阶段的行动对应不同的策略性行动，我们把策略性行动分为三类：承诺、威胁及许诺。这三个策略性行动的目的都是为了改变第二阶段博弈的结果从而对自己有利。它们三个究竟哪个能达到目的是依具体的情形而定的。但是，最重要的是，你必须让其他的参与人相信你在第二阶段确实会做你在第一阶段中宣布的事。换句话说，策略性行动是否具有置信度尚不知晓，而只有可信的策略性行动才会产生预期的效果，就像第 9 章中所举的例子一样，仅仅宣布是不够的。在第一阶段，你必须做一些辅助工作让人相信你会在第二阶段做已宣布的事。我们将研究会给你带来好处的两类两阶段博弈以及

第一阶段中进行的担保行动。

你对策略性行动的运用及置信度的熟悉可能会超乎你自己的想象。例如，父母常使用威胁（吃完青菜才能吃点心）或许诺（如果你的成绩能维持在B水平，期末时就送你一辆脚踏车）来影响小孩的行为。许多小孩都知道这些威胁或许诺大多不可信，因为当小孩做坏事的时候，他可以撒娇说下次不会再犯来逃避“威胁”时提出的处罚，虽然小孩自己的许诺也并不一定可信。而且，小孩长大后会变得更在意自己的外表，这时他们会就运动及饮食问题对自己做出承诺，虽然这些承诺也并不一定可信。所有的这些方法——承诺、威胁和许诺——只是策略性行动的例子。它们共同的目的是在博弈的下一阶段改变其他参与人的行动，甚至是自己未来的行动。但是除非这些策略性行动是可信的，否则是不能达到目的的。在本章，我们将会使用博弈理论来系统地研究如何使用这些策略以及如何使它们可信。

然而，需要提醒的是使其具有置信度既困难且微妙。这里将针对策略性行动如何起作用提供一些基本原则和总体上的理解，即所谓“策略科学”。但是实际运用中你需要对博弈情景有很好的理解，才能使策略性行动有效，否则你的对手会因为比你能更好地理解博弈概念或情景而占你的便宜。所以，在实践中运用策略性行动是含有艺术成分的。它也存在风险，尤其是当你使用**边缘政策**（brinkmanship）时，有时候会导致灾难。你可以成功地将这些思想付诸行动并且毫发无损，但是一定要记住我们的忠告：用这些策略时别让他人承担风险。

10.1 策略性行动的分类

因为策略性行动的使用十分依赖行动的次序，所以研究它们时，必须知道“先行动”的含义是什么。过去我们认为这个概念无须详细说明，但现在我们要更准确地了解它。“先行动”包含了两个方面的含义：首先，你的行动对于别人来说必须是**可观察的**（observable）；其次，它是**不可逆的**（irreversible）。

这里只考虑两个参与人——A与B——之间的策略性行动。A先行动，如果A的选择没有被B观察，那么B就不能对其做出反应，二者的行动也就不相干。例如，假设A与B在拍卖会中喊价。A在星期一秘密决定其竞标价格，B在星期二决定，两个价格分别被寄给拍卖者，并在星期五公开。当B做决定时，它不知道A所做的决定，所以它们的行动在策略意义上可以看成是同时的。

如果A的行动是不可逆的，那么A可能会假装做一件事来诱使B做出反应，然后再改变自己的行动来获利。当然，B应该预料到A的圈套，不应该被骗，所以它不会对A的选择做出反应。所以，在真正的策略意义上A并没有先行动。

对可观察性及不可逆性的考虑会改变策略性行动的本质、类别及策略性行动的置信度，这里从策略性行动的分类开始。

10.1.1 无条件的策略性行动

这里假设 A 让他的行动在第一阶段是可观察的和不可逆的。他会对外宣布："在接下来的博弈中，我会做出行动 X。"这表示 A 未来的行动是无条件的，不管 B 做什么 A 都会做 X。如果可信的话，这相当于改变了第二阶段的次序，使 A 先行动 B 后行动，A 的行动是 X。这种策略性行动称为**承诺**（commitment）。

如果博弈本来的规则就是在第二阶段 A 先行动，那么这个宣布是可信的；但假如在博弈的第二阶段 A、B 同时行动或 A 后行动，那么假如这个宣布是可信的，它就会改变结果，因为它改变了 B 对行动结果的看法。这样的承诺会很容易让 A 获得充当先行动者的优势。

在第 3 章中的街心花园博弈中，三位女士进行了一场序贯行动博弈。在这场博弈中，每个人必须决定是否为修建公共花园作贡献，两个或两个以上的人作贡献才会修一座像样的花园。逆推均衡是先行动者（爱弥莉）选择不作贡献而其他参与人（莱娜和塔尼娅）则选择作贡献。然而，通过做出一个不作贡献的可信的承诺，塔尼娅（或莱娜）可能会改变博弈的结果。即使在爱弥莉和莱娜已经将她们的决定公布于众之后，塔尼娅才公布她的决定。塔尼娅可让大家都知道她已经将所有的积蓄（和精力）用在一所大房子的修葺工程上了，这样她就绝对没有剩余的钱用在花园的建设上了。那么在爱弥莉和莱娜做出决定之前，塔尼娅就已经肯定了她不会作贡献。换句话说，塔尼娅改变了整场博弈，使自己成为实际上的先行动者。你能发现新的逆推均衡是爱弥莉和莱娜对花园的建设作出了贡献，她们二人的均衡支付是 3，而塔尼娅的是 4——均衡的结果与塔尼娅是先行动者有关。在下面的各节中将会给出更多的关于承诺的例子。

10.1.2 有条件的策略性行动

另外，A 有可能在第一阶段宣称："在接下来的博弈中，我会对你的选择做出反应。如果你选择 Y_1，我将会做 Z_1；如果你选择 Y_2，我将会做 Z_2……"换句话说，A 的行动是以 B 的行动为条件的，我们称这样的行动是**反应规则**（response rule）或反应函数。A 的意思是说，在博弈的第二阶段，他会后行动，并会像在第一阶段中所宣布的那样对 B 的选择做出反应。为了使这样的宣布有意义，A 必须在第二阶段一直等待，直到他观察到了 B 采取了不可逆转的行动才采取行动。换句话说，在第二阶段 B 要成为真正意义上的先行动者。

有条件的策略性行动有许多不同的形式，依参与人的目标及实现目标的方式而定。当 A 想阻止 B 做某件事时，我们说 A 试图遏止 B 或达成**遏止**（deterrence）；当 A 想引导 B 做某件事时，我们说 A 试图强迫 B 或达成**强迫**（compellence）。在后面我们会再讨论一下它们的区别以更加明确这些概念的含义。接下来我们看看实现这些目标的方法。如果 A 宣称："除非你的行动（或者并非行

动，可能是事件）符合我已表达的意愿，否则我将伤害你。”这就是一种**威胁**（threat）。如果A宣称：“如果你的行动（或者并非行动，可能是事件）符合我的意愿，我就会奖励你。”这是一种**许诺**（promise）。此处的“伤害”和“奖励”是以支付衡量的——当A伤害B，A会使B的支付降低；当A奖励B，A会使B的支付提高。威胁和许诺是我们重点分析的两种有条件的策略性行动。

为了了解这些策略的本质，需要考虑之前提到的晚餐博弈。在一般的按时间顺序的行动中，首先小孩决定要不要吃青菜，然后父母决定给不给小孩点心，逆推法告诉了我们结果，小孩拒绝吃青菜，因为他们知道父母不会让他们挨饿和不高兴，而会给他们点心。然而父母预料到这个结果后，就会首先采取行动以改变这一结果，也就是宣布一个“除非吃青菜，否则没有点心”这种形式的条件反应规则，这个宣布就是一个威胁。这是博弈前的第一步行动，它决定了在接下来的实际博弈中小孩将如何采取第二步行动。如果小孩相信这个威胁，就会改变逆推计算，即去掉博弈树的一个分支——即使小孩没吃青菜，父母也会给点心。这可能会改变小孩的行为；父母也希望这样就会使小孩照他们的意愿去做。同理，在学习博弈中，给予新脚踏车的许诺会使小孩学习更努力。

10.2 策略性行动的置信度

我们已经看到，A的策略性行动会改变其他参与人的支付，那么A的支付会怎样呢？如果B的行动符合A的意愿的话，A就会得到更高的支付，但是A的支付也会受他自己反应的影响。在威胁的例子中，如果B不照A的意愿行动，则A的威胁的实施反而会影响到自己的支付：父母看到小孩因没有吃到点心而不高兴，自己也会心疼。同理，在许诺的例子中，如果B照A的意愿行动，则A给B的奖励会影响到自己的支付：小孩用功读书的话，父母必须花钱买奖品，但他们乐于看到小孩快乐的样子及优异的成绩。

A的支付效应会严重地影响他的策略性行动的有效性。以威胁为例。如果A的支付由于威胁行动的实施而增加的话，则B就会推论不论自己是否照A的意愿行动，A都会实施威胁；所以B就没有照A意愿行动的动机，这个威胁就是无效的。例如，如果父母是虐待狂，乐于见到小孩没吃到点心，则小孩就会想：“反正我是吃不到点心了，为什么要吃青菜呢？”

所以威胁最重要的地方就是威胁者必须为自己实施威胁行动付出代价。在晚餐博弈中，父母必须乐于给小孩点心。在真正的策略意义上的威胁具有让被威胁者付出代价的固有特性，这就是互相伤害的威胁。

从技术的角度看，威胁决定了你在接下来博弈中的策略（反应规则）。一个策略必须明确说明每个博弈树分支的最后结果。所以“如果你不吃青菜就没有点心”是一种不完整的策略描述，应该补充上“如果吃青菜就有点心”。但威胁通常不指第二部分，为什么呢？因为第二部分是暗含的，自动成立。为了让威胁有效，策略的第二部分——暗含的许诺——也应该是自动可信的。

所以“不吃青菜就没有点心”的威胁中包含着“吃青菜就有点心”的许诺，且只有当许诺可信时威胁才有效。在这个例子中，只要父母乐于见到小孩吃点心，暗含的许诺自动具有置信度。换句话说，因为威胁行动的付诸实施会让父母付出代价，所以暗含的许诺自动可信。

从另一个方面来看，威胁规定：如果你的意愿没达成，你将会做一些事；如果这种情况真的出现，你将会遗憾地不得不去做这些事。那么为什么在第一阶段又要做出这些规定呢？为什么要选择束缚自己呢？拥有选择的自由不是更好吗？在博弈理论中，拥有更多的选择不一定是好事。至于威胁，在第二阶段失去自由反而是有其策略价值的。它改变了你的对手对你未来反应的预料，而这样的改变将对你有利。

许诺也有类似的效应。如果小孩知道父母乐于给他礼物，他将会预料在不久的将来的某些场合他总能得到脚踏车，例如在即将到来的生日。那时买脚踏车的许诺对促进小孩努力学习就没什么作用了。为了使策略有效，许诺的奖励是要付出一定代价的，让另一参与人不会觉得他总能得到这一奖励（这在策略上是很重要的一点，你应该提醒你父母他们许诺的奖励应该更大或更昂贵，而不只是让你高兴的小礼物）。

无条件的策略性行动也是一样（承诺也是一样）。在讨价还价的例子中，别人知道你有行动和谈条件的自由，所以“不退让”的承诺是你较好的打算。如果你提出要60%的“馅饼”，而对方只给你55%，你也许会接受（此时你受到“引诱”，背离了你的初衷）。但是如果你事前宣称你将不接受小于60%的“馅饼”且极具可信度，则“引诱”将不会出现，你接受只有55%的馅饼这样的情况就不会发生了，而你将会得到较好的结果。

这正是策略性行动的本质——当第二阶段的博弈确实要求你实施先前计划好的一切时，你又极不乐意。这对所有类型的策略性行动都适用，也是使置信度成为一个问题的地方。你必须在第一阶段做一些事来确立你的置信度——让其他参与人相信他们的行动如违你所愿，你一定会将计划付诸实施，而不会因利益的“引诱”而放弃，这样做是为了使你的策略性行动有效。这就是为什么放弃你行动的自由对你是有利的。另外，你可以通过改变自己在第二阶段的支付来获取威胁或许诺的置信度，方法是按自己在第一阶段宣称的那样去行动从而得到最优结果。

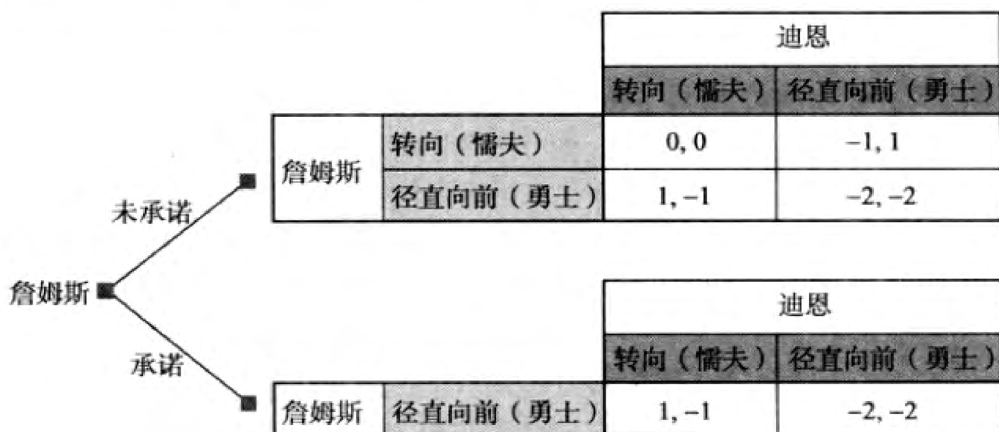
因此，有两种方法使你的策略性行动可信：(1) 在你未来的选择组合中排除“引诱行动”；(2) 减少这些“引诱行动”的支付，让计划的行动成为真正最佳行动。在接下来的几节中，我们会首先阐明策略性行动的机制（假定这些策略性行动是可信的）；接着对置信度进行评论，并把对置信度的一般分析留在本章的最后一节。

10.3 承　　诺

我们在第4章研究过懦夫博弈，并找到了两个纯粹策略的纳什均衡。每个参

与人都选择自己“径直向前”而另一参与人“转向”的均衡。[1]在第 6 章我们可以看到，如果博弈行动是序贯的而不是同时的，先行动者会选择“径直向前”，后行动者会选择“转向”（而不是相撞）。现在我们从另一个角度来看相同的问题。即使博弈行动是同时的，如果一个参与人进行了策略性行动——创造了第一阶段，并使人们相信他会在懦夫博弈即第二阶段中实施他在第一阶段宣称的行动，他将通过承诺“会采取强硬行动即‘径直向前’”得到与先行动者一样的好处。

虽然观点简单，但这正是正式分析（深入理解及发展技巧以解决复杂问题）的提要。这里的两个参与人是詹姆斯和迪恩。假设詹姆斯是有机会采取策略性行动的一位。图 10—1 展示的是两阶段博弈树。在第一阶段，詹姆斯不得不决定是否做出承诺。沿着上分支，他不做出承诺，在第二阶段是同时行动博弈，它的支付与表 4—14 和图 6—5 所示的一样。这个两阶段博弈有多个均衡，詹姆斯只在其中一个均衡中拥有最佳支付。沿着下分支，詹姆斯做出承诺，在这里我们解释为詹姆斯将放弃行动的自由，即他在第二阶段只有“径直向前”这一可行的行动。此时的第二阶段博弈表对于詹姆斯只有一行（对应于他宣布的“径直向前”的行动）。在这个表中，迪恩的最佳行动是“转向”，所以均衡的结果是詹姆斯得到了最佳支付 。因此，在第一阶段，詹姆斯发现做出承诺是最优的，这一策略性行动会确保他得到最佳支付，然而，如果不做出承诺则结果不确定。



		迪恩	
		转向（懦夫）	径直向前（勇士）
詹姆斯	转向（懦夫）	0, 0	−1, 1
	径直向前（勇士）	1, −1	−2, −2

		迪恩	
		转向（懦夫）	径直向前（勇士）
詹姆斯	径直向前（勇士）	1, −1	−2, −2

图 10—1　懦夫博弈：限制行动自由的承诺

詹姆斯如何使承诺变得可信呢？就像任何最先的行动一样，承诺行动必须是（1）不可逆的，（2）对于其他参与人是可观察的。人们提出过一些极端和有趣的建议：詹姆斯可以拆掉方向盘，把方向盘扔到窗外，这样迪恩就看到詹姆斯不能够转弯了；詹姆斯也可以把方向盘固定住，使其不再能转动，但是要想向迪恩证明方向盘确实被固定了而且不易被松开是不容易的。这些方法轻松地将转向这一选择从詹姆斯的第二阶段的可行选择中排除掉了，让他只能选择“径直向前”。

似乎更言之成理的是，如果此博弈每周末都进行的话，詹姆斯会得到强硬的名声，这就成了他之后任一天行动的保证。换句话说，詹姆斯可以通过“转向”来改变自己的支付，此时，其声誉就会受损，从而其支付将减少。如果减少数量足够大的话，例如 3，在詹姆斯做出承诺后的第二阶段的博弈就会有一个不同的支付表。这个博弈的完整的博弈树见图 10—2。

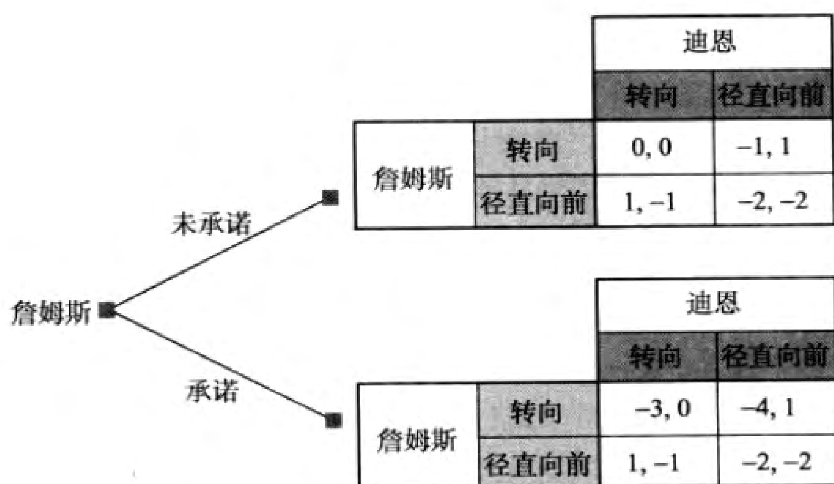


图 10—2　懦夫博弈：改变支付的承诺

现在，在做出承诺后的第二阶段，"径直向前"就变成了詹姆斯的最佳选择。事实上，这是他在这个阶段的占优策略。迪恩的最佳策略是"转向"。在第一阶段预料到这一结果后，詹姆斯发现，如果做出承诺他可以得到 1（改变了第二阶段的支付），如果不做出承诺，他不一定可以得到支付 1，也许支付更少。逆推法的结论就是詹姆斯应该做出承诺。

两个参与人（或所有参与人）都可以做出承诺，所以成功与否取决于你抓住先行动时机的速度及你所进行策略性行动的置信度。如果存在观察上的时滞，双方可能同时做出承诺：每个人都跟别人同时拆掉自己的方向盘，然后扔到窗外，那么车祸便不可避免了。

即使一个参与人有做出承诺的优势，其他参与人也可以挫败先行动者进行承诺的企图。他可以表明没有"看见"对方承诺的能力，例如断绝交流渠道。

懦夫博弈可能已经不合时宜了，但是第二个例子应该是反复出现而为大家所熟悉的。在班上，老师对交作业的期限的态度可以是"强硬"或"软弱"，学生可以"准时交"或"迟交"。表 10—1 展示了该博弈的策略式。老师们并不希望自己铁石心肠，因此他们最好的结果（支付是 4）是当自己态度软弱时，学生仍准时交作业；最不好的结果就是他的态度强硬，学生仍是迟交。而处于中间的两个策略，他承认准时的重要性，认为（强硬，准时）要好于（软弱、迟交）。学生最喜欢（软弱，迟交）的结果，因为他们可以整个周末狂欢也不会遭到任何迟交作业的惩罚。与老师一样，（强硬，迟交）是他们最不愿意看到的结果。在剩下的两个策略中，他们宁愿（软弱，准时交）也不愿意（强硬，准时交），因为假

表 10—1　　交作业期限博弈的支付表

		学生	
		准时交	迟交
老师	软弱	4，3	2，4
	强硬	3，2	1，1

如他们考虑到准时交作业是出于自己的决心而不是由于受到会被惩罚的威胁，这时他们会比较有自尊。[2]

如果这是一个同时行动博弈，或者是老师后行动，则"软弱"就是老师的占优策略了，而学生就会选择"迟交"。均衡的结果是（软弱，迟交），支付是（2，4）。但是如果老师从一开始就承诺"强硬"的话可能会达到一个较好的结果。我们这里不像图 10—1 和图 10—2 一样再画一棵博弈树了。这棵博弈树和懦夫博弈非常相似，这里留给你们自己画。当老师不做出承诺时，第二阶段就和以前一样，老师得到支付 2；当老师承诺"强硬"时，学生发现在第二阶段准时交会更好，则老师会得到支付 3。

老师的承诺行动不同于他在同时行动博弈中采取的行动，事实上承诺行动是学生行动之后老师采取的最优行动。这就是需要使用策略思维的地方。老师若是宣告"软弱"就得不到什么利益；而学生则希望老师什么承诺都不做。为了从策略性行动中获得好处，他所做的承诺不能跟同时行动博弈的均衡策略一样。这个策略性行动改变了学生的期望和行动。一旦他们相信老师的承诺，他们将会选择准时交作业。如果他们尝试迟交一次，老师可能会原谅他们，理由是"只此一次"。但这里存在着使策略性行动不付诸实施的诱惑，使承诺的置信度下降。

更为戏剧性的是，在这一例子中，使老师获益的策略性行动是采取劣策略。他承诺"强硬"，这相对于"软弱"来说是劣势的。如果学生准时交，"强硬"会使老师得到 3；如果学生选择"迟交"，则老师选择"强硬"只能使其得到 1。然而如果老师选择"软弱"，对应的支付分别是 4 和 2。你是否觉得采取劣策略倒可以获益是件不可思议的事情？这里就必须扩展占优的概念，使其适用范围扩大。占优包含以下两种计算方法：（1）在对方行动后，我做何反应？什么是最好（或最坏）的选择？（2）如果是同时行动，对方采取行动 X，那么究竟什么对我来说是最好的（或最坏的）策略呢？对所有的 X 行动，我的最好（最坏）策略都一样吗？当你先行动时，这两种计算方式都不具有实际价值；相反你必须预想到他人的反应。所以老师并不比较支付表中垂直相邻的支付（一次只考虑学生的一种可能的行动）。相反地，他会计算学生对他每一行动的反应。如果他承诺"强硬"，他们就会选择"准时交"，但是，如果他承诺"软弱"（或未承诺），他们就会选择"迟交"。所以唯一恰当的是将支付表的右上格和左下格进行比较，老师会选择左下格。

老师所做的承诺必须符合先行动的条件，这样的承诺才可信。首先，它必须在其他参与人之前行动。老师必须在交作业之前制定有关交作业最后期限的强制规则。接下来，它必须是可观察的——学生必须知道他们应该遵守的规则。最后，可能也是最重要的，它必须是不可逆的——学生必须知道老师不管在什么情况下都不会改变主意或者是原谅他们。好心的老师只会让学生找借口迟交，并说"下次不要再犯"之类无关痛痒的话。

老师也可以为了建立置信度拿学校的规定做挡箭牌。只要在第二阶段从他可行的选择组合中剔除"软弱"即可，或者和懦夫博弈一样，他可以建立强硬的名声，并且会为失去名声付出较高的代价，这样他就可以改变自己的支付了。

10.4 威胁与许诺

这里要强调的是，威胁和许诺是反应规则：你未来的实际行动会依照其他参与人的行动而定，但是你未来行动的自由会受反应规则的限制。再次重申，这样做的目的是改变其他参与人的期望，使他们的行动对你有利。在这个过程中最重要的是，你必须遵守这个规则，如果你在后来的行动是自由的，你就不会想去遵守这个规则。我们会再一次地阐明策略性行动取得置信度的原则，但仍要提醒你的是，实际的执行仍然是一门艺术。

记住第 1 节的分类。如果其他参与人不依照你的意愿行动的话，“威胁”这条反应规则将对其他参与人产生一种不好的后果。如果其他参与人依照你的意愿行动，“许诺”这条反应规则将对其他参与人产生好的结果。每一种反应或者可阻止其他参与人做某事，或者可强迫其他参与人做某事，我们会依次指出它们各自的特征。

10.4.1 威胁的例子：美日贸易关系

本例来自于美日两国长期的贸易摩擦。每个国家都有权利选择对另一国的货物开放本国市场或是关闭本国市场。对于结果，两国有不同的偏好。表 10—2 显示了贸易博弈的支付表。对于美国，最好的结果（支付是 4）是两国市场都开放，部分原因是它对市场系统和自由贸易的承诺，部分原因是和日本做贸易所获得的利益——美国消费者可以获得高品质的汽车和电子产品，并且可以出口他们的农产品和高技术产品。类似地，最坏的结果（支付是 1）是两国市场都关闭。在只有一国市场开放的两个结果中，美国宁愿开放本国市场，因为日本市场相对较小，失去出口市场的损失要比失去进口本田汽车和随身听的损失小。

表 10—2　　美日贸易博弈的支付表

		日本	
		开放	关闭
美国	开放	4，3	3，4
	关闭	2，1	1，2

对于日本，在这个例子中，我们认为它是保护主义者，并且是生产导向型的。它的最好的结果是美国市场开放而本国市场关闭，最坏的结果则刚好相反。至于另外两个结果，它宁愿两国市场都开放，因为它的产品可以进入庞大的美国市场。[3] 双方都有优势策略。不管是同时行动博弈还是序贯行动博弈，均衡的结果是（开放，关闭），支付是（3，4）。这结果正好符合美国人对两国贸易政策的

一般印象。

在此均衡中，日本获得了最佳支付，所以没有必要做任何策略性行动。然而，美国会试图用支付 4 代替 3。但是，在此例中一个普通的无条件承诺将没什么作用。不管美国做出什么样的承诺，日本的最佳反应都是关闭本国市场。而承诺开放本国市场对美国较为有利，这是不需要进行任何策略性行动都可以达到的均衡。

但假设美国选择如下的条件反应规则："如果你关闭你方市场，我们也将关闭我国市场。"这时就会变成图 10—3 显示的两阶段博弈。如果美国不使用威胁的话，第二阶段就和以前一样，会造成美国市场开放，得到的支付为 3，而日本市场关闭，得到支付 4。如果美国使用了威胁，则在第二阶段只有日本有选择的自由，不管日本做何选择，美国都会根据反应规则行动。所以，沿着博弈树的分支，只有日本是积极的参与人，现在写下两者的支付：如果日本市场关闭，美国市场就关闭，美国会得到支付 1，而日本得到支付 2。如果日本开放市场，则美国的威胁就发生了作用，它也乐意开放本国市场，得到支付 4，而日本得到支付 3。对于这两种可能性，第二种对于日本是更好的选择。

现在使用我们熟悉的逆推法来分析：美国在知道第二阶段所有最终结果后，在第一阶段使用威胁的方法会对自己较有利。威胁会使日本开放它的市场，而美国能得到最好的结果。

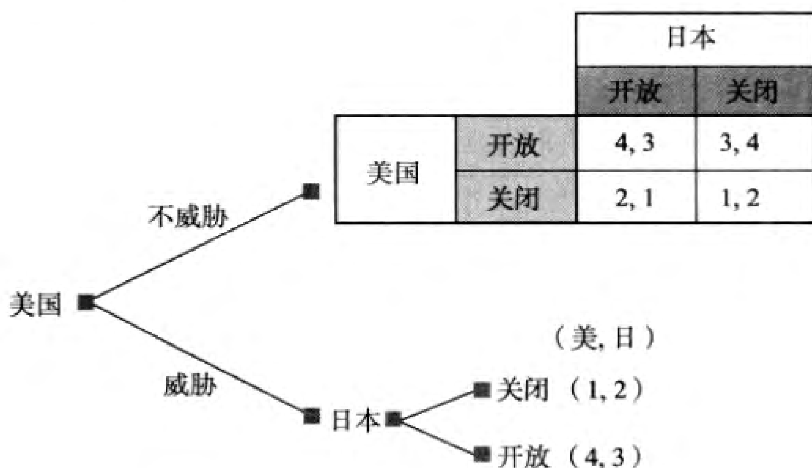


图 10—3 考虑威胁的美日贸易博弈树

描述过威胁的机制后，现在指出它的一些重要特征。

(1) 当美国使用可信的威胁时，日本将不采取占优策略——"关闭"。再者，占优的概念也只是跟同时行动或日本后行动的情形有关。这里，日本知道美国会采取偏离其占优策略的行动。在支付表中，日本将在左上格或右下格中做出选择，它会选择右下格。

(2) 要建立威胁的置信度是一件麻烦事，因为如果日本试探着关闭本国市场，美国就面临着是否实施威胁的考验。事实上，如果威胁行动（或者说实施威胁）是美国应对日本关闭其市场的最佳行动，那美国也就没有必要事先威胁了（但美国还是要发出警告让日本知道这种情形）。策略性行动是为了让参与人做某

件事，而不是真正希望在对方选择不做该事时按策略性行动所规定的那样去做。就像早先解释的那样，威胁对于威胁者来说必须付出代价，也就是威胁行动会互相伤害。

(3) 条件规则“如果你关闭市场我也关闭市场”并不是对美国策略的完整描述。为了使其完整，应该另外加一句来说明日本市场开放后美国的反应：“如果你开放市场我也开放市场。”这附加的一句是暗含的许诺，也是威胁的一部分，但它不需要明显地说出来，因为它是自动可信的。给定第二阶段的支付，如果日本市场是开放的，对于美国来说开放本国市场就是最佳选择。如果情况并非如此，即使日本市场开放，美国仍关闭本国市场，则暗含的许诺必须明白地说出来并且以某种方式使之具有可信性。否则，美国的威胁就等同于无条件承诺：“我们将关闭我国市场。”这样就不可能诱使日本做出美国希望的反应。

(4) 如果威胁可信的话，它会改变日本的行动。依照最初的状况，我们可以把这看成遏制或者强迫。如果日本市场一开始就是开放的，且日本正考虑向保护主义转变，而美国的威胁会遏制这种行动。但是，如果日本市场一开始就是关闭的，美国的威胁就会强迫其开放市场。所以策略性行动是遏制还是强迫要取决于现状。这两者好像只有意义上的区别，但是在实际中策略性行动的运作及置信度却深受此区别的影响，这个问题在本章后面会进一步讨论到。

(5) 这里美国有一些可使其威胁可信的方法。第一，立法在适当条件下实施威胁。这会减少第二阶段可行选择集合中的引诱行动。世界贸易组织的一些关系条款有此作用，但过程是缓慢且不确定的。第二，将任务交由像美国商业部这样的机构执行。商业部由美国生产者掌控，他们希望保持市场关闭以减少竞争压力。这会改变美国第二阶段的支付。用商业部的支付代替真正的美国的支付，其结果是威胁行动成为最佳行动（这样做的风险是即使日本市场开放，商业部仍坚持保护主义立场，所以威胁的可信度的建立可能导致暗含的许诺的置信度的丧失。）

(6) 如果威胁起作用了，它并不一定要执行。这样的代价才是无足轻重的，事实上，当威胁行动有错估的风险或错误实施（当对方已经如你所愿地行动了，却仍然实施威胁）的风险时，让我们避免使用威胁。更明白地说，如果日本不买美国的稻米或半导体，美国就威胁退出防御联盟，这个威胁的代价太大、风险太高，美国不会去执行，所以并不可信。

并不是所有的威胁都应采纳，一方面威胁既要大到让对方惧怕它，而按照自己的意愿去做，另一方面，威胁也不能有太大的风险，使先行动者都不可能实施，从而失去置信度。如果威胁太大不便于实施，就通过随机实施来降低威胁。美国对日本不能说“如果你不开放你的市场，我们就会拒绝保护你”，而应该说“如果你不开放你的市场，我们双方的关系会恶化，我们国会可能会拒绝在你受攻击时提供帮助，即使我们建立了防御联盟”。事实上，美国可以刻意制造一种情绪，以提高国会实施威胁的概率，这样日本就会明显地感到威胁了。这种威胁，带有风险，但不一定会带来不好的结果，被称为边缘政策。这是相当微妙和危险的策略性行动，我们将在 14 章具体讨论这个问题。

(7) 日本在美国使用威胁这一策略性行动后得到的结果比在美国不使用威胁

时得到的结果差，所以它也会采取策略性行动来阻止美国使用威胁。例如，假设日本市场目前关闭，美国试图强迫它开放。日本可以原则上同意开放，实际敷衍，恳请必要的延缓开放，因为需要时间去集中必要的政治上的共识从而制定市场开放的法律，然后再恳请延缓开放，因为需要补充必要的管理规则，诸如此类。因为美国不想率先实施威胁，因此每一次都存在接受延缓开放的引诱。或者日本宣称它的外交政策使完全开放市场很困难，希望美国能接受日本对少部分产业的保护。日本的借口和提出的要求会越来越多，而每一次（每个小步骤）都不足以让美国实施威胁从而导致贸易战争的爆发。这种以小步骤或切片方式使强迫变得无效的方法，称为“腊肠策略”（salami tactics）。

10.4.2 许诺的例子：餐馆定价博弈

现在我们使用第 5 章中的餐馆定价博弈来描述许诺。我们看到第 5 章中的那个博弈是个囚徒困境，在这里，为简单起见，我们假设只有两种价格选择：26 美元的共同最佳价格和 20 美元的纳什均衡价格。这个版本的博弈中的每家餐馆的利润可以通过 5.1 节中的函数计算出来。结果见表 10—3。没有进行任何策略性行动时，博弈均衡是双方都采用 20 美元的低价，但采用高价 26 美元会有更高的支付。

若某一方做出可信的许诺：“如果你采取高价，我也会采取高价”，则二者将获得合作的结果。例如，如果泽维尔做出了该许诺，则伊冯娜会知道选择 26 美元的价格将达成双赢，产生表中右下格的支付；而选择 20 美元会使泽维尔同之前一样也选择 20 美元的价格，得到左上格的支付。在这两者之间，伊冯娜更愿意选择前者，所以会选择高价。

表 10—3　旅馆囚犯困境的支付（百美元/月）

		伊冯娜的炭火西餐馆	
		20（低）	26（高）
泽维尔的达帕斯西班牙餐馆	20（低）	288，288	360，216
	26（高）	216，360	324，324

通过画出这个两阶段博弈的博弈树可以更好地进行分析，其中，泽维尔在第一阶段需要做出是否做出许诺的选择。这里省略博弈树，部分原因是你可以在画这棵博弈树的过程中加深对这场博弈的理解，另外是想说明熟悉这些观点后有些细节就变得无关紧要了。

泽维尔的许诺是否可信值得怀疑。为了对伊冯娜的行动做出反应，泽维尔必须在第二阶段后行动。相应地，伊冯娜在第二阶段要先行。记住，先行是一种不可逆并且可观察的行动。所以如果伊冯娜先行动选择高价，就容易受到泽维尔的欺骗。当泽维尔知道伊冯娜处于易受欺骗的地位，它很可能受到引诱违反许诺。所以泽维尔必须让伊冯娜相信，当它选择高价时自己不会被引诱而选择低价。

怎么办呢？也许泽维尔的老板可以把价格的决定权下放给地方经营者，并白纸黑字写下指示：如果伊冯娜采取高价，它也会采取高价，泽维尔的老板可以邀请伊冯娜的老板目睹整个过程，然后远航旅行不再干涉此事（即使如此，伊冯娜的管理层还是会怀疑，泽维尔可能秘密通过电话或上网进行干涉）。这就等于从泽维尔的第二阶段选择中排除了欺骗行为。

或者泽维尔可以更广泛地发展在生意上和社区中遵守信用的声望。在重复关系中，许诺会发挥作用，因为违反许诺有可能会破坏未来的合作。本质上，持续的关系表示把博弈分成许多小部分，在每一部分中违反许诺所带来的获益远小于它的损失。在这类博弈中，未来合作失败的代价会改变欺骗得来的支付。[4]

我们在前面说过，每个威胁都有暗含的许诺。同理，每个许诺也都暗含着威胁。在这个例子中，暗含的威胁是："如果你采取低价我就采取低价。"它不需要明确地说出来，因为它自动可信——它描述了当伊冯娜采取低价时泽维尔的最佳反应。

当然，威胁和许诺有一个很重要的不同。如果威胁成功了的话，它就不必实施了，威胁者就不用付出代价。所以威胁可以比所需的更大，这样的威胁才更有效（虽然威胁太大会有危险，甚至可能会失去置信度）。如果许诺成功地如自己所愿改变了他人的行动，那么许诺人必须做他已经许诺的事情——当然这是要付出代价的。在前面的例子中，代价仅是放弃欺骗的机会，但获得了最高支付，在其他情况下，许诺者提供实际的礼物给他人，其代价是更真实的。无论在何种情况下，许诺人都力图把这种代价减小——代价只需大到使许诺有效即可。

10.4.3 威胁和许诺相结合的例子——中美政治行动

当我们分别考虑威胁和许诺时，明确的威胁包括了自动可信的暗含的许诺，反之亦然。然而有些情况下两者的置信度都有问题，那么策略性行动就要让两者都明确且可信。

明显的威胁和许诺相结合的例子来自于多个国家合作处理邻国危险局势。具体来说，我们以中美两国是否应该考虑采取行动使朝鲜放弃其核武器计划为例，在表 10—4 中给出了中美两国分别在选择“行动”或“不行动”时的支付表。

表 10—4　　中美政治行动博弈的支付表

		中国	
		行动	不行动
美国	行动	3，3	2，4
	不行动	4，1	1，2

每个国家都希望另一个独自采取行动来反对朝鲜核武器计划，所以右上格对于美国来说有最佳支付 4，左下格对中国来说有最佳支付 4。对于美国来说最坏的情况就是都不采取行动，那样的话核战争的威胁会增大，美国是不能接受的。

对于中国而言，最坏的结果是独自采取行动，因为会有巨额花销。两者都认为联合行动是次好的（支付是 3）。美国在单独行动的情况下可以得到支付 2。而对于中国，都不行动可以得到支付 2。

若不进行任何策略性行动，该干涉博弈是占优可解的。不行动是中国的占优策略，行动是美国的最佳选择。均衡的结果就是右上角，美国得到支付 2，中国得到支付 4。中国得到了最好的结果，所以没有任何理由采取策略性行动。但是美国可以做一些事使其得到的支付不止是 2。

美国做何种策略性行动会提高其均衡支付呢？无条件的行动（承诺）是没用的，因为中国会用"不行动"来回应美国的任何先行动。单独的威胁（如果你们不行动，我们就不行动）也没用，因为暗含的许诺（如果你们行动，我们也行动）是不可信的——如果中国行动，美国会撤退，把一切事情留给中国，它得到的支付是 4 而不是兑现许诺时所得到的支付 3 了。只是许诺也没用，因为中国知道如果中国不干涉美国也会干涉，美国做出的"如果你们干涉，我们就干涉"的许诺成为了一个简单的干涉承诺了，中国可以置身事外，得到最高支付 4。

在这个博弈中，美国发出的明确的许诺必须有暗含的威胁："如果你们不干涉，我们也不干涉"，但是威胁又不是自动可信的。同理，美国明确的威胁也必须有暗含的许诺："如果你干涉，我们也干涉。"但是，那也不是自动可信的。所以美国必须做出明确的威胁和明确的许诺。它必须把威胁和许诺联合起来："只有你们干涉了，我们才干涉。"这需要威胁和许诺都是可信的。通常这需要订立涵盖整个关系的条约才能建立这种置信度，而并不是当情况发生时才订立临时的谈判协议。

10.5 一些额外的议题

10.5.1 策略性行动何时起作用？

我们已经从例子中看到，和没有采取策略性行动的原始博弈相比，策略性行动可以给一方或另一方带来更好的结果。而什么时候需要这种策略性行动呢？

无条件行动——承诺——并非每次都会为做出承诺者带来优势。事实上，如果原博弈存在后动优势，那么参与人首先做出承诺就是个错误，因为他成了先行动者。

至于条件策略性行动——威胁和许诺——就不可能带来坏处。最坏的结果跟没有制定条件反应规则时一样。然而，如果条件策略性行动带来了利益，那一定是因为条件反应规则使某个行动代替了原有最佳行动，成为了新的最佳行动。因此，当威胁和许诺能带来正的效应时，它们的置信度就存在问题了，必须采取相应措施来建立置信度。我们在前面几个例子中谈过这些措施，在后面将更广泛地讨论如何获得置信度。

那么什么情况下需要成为策略性行动的接受者？没有人想被威胁，如果知道

对方会使用威胁，你可以通过寻找另一种先行动来获利——这一行动使威胁的置信度和威胁性降低。我们会稍微分析一下这样的例子。然而，让别人对你做出许诺常常是必要的。事实上，当一方做出可信的许诺时，双方都可能获益，就像本章所讲的餐馆定价博弈（囚徒困境）一样，如果许诺达到了双方合作的目的，那么一方或者双方做出许诺可能会促进双方收益的提高。

10.5.2 遏制对强迫

原则上，不论是威胁还是许诺，都可达到遏制或强迫的目的。例如，一对希望小孩努力学习（强迫）的父母会做出许诺在小孩取得优异成绩时给予奖励（一辆新的脚踏车），也可以做出威胁如果小孩没有取得足够好的成绩就进行惩罚（下学期晚上不能出去玩）。同理，想要小孩远离坏同伴的父母（遏制）既可以采取奖励（许诺）措施，也可以采取惩罚（威胁）措施。实际上，这两种类型的策略性行动起作用的方式不同，将影响各自的使用范围。一般来说，遏制最好使用威胁方式，强迫则最好使用许诺方式，原因在于时机与动机的不同。

遏制威胁可以是消极的——只要对方不去做你试图遏制的事，你就不需要做任何事。它也可以是静态的——你不需要附加任何时间限制，因此你设立威胁条件，然后把其他的事交给对方就可以了。所以希望小孩远离坏同伴的父母说："如果让我知道你又和 X 一起的话，以后整整一年中晚上 7 点钟后都不能出去玩。"然后父母就等着瞧，只有当小孩的行为违反了父母的意愿时，父母才实施威胁。通过许诺来达到相同的遏制就要求更复杂的监控和连续的行动："在每个月的月末，如果我确定你没和 X 一起的话，就会给你 25 美元。"

强迫必须有截止日期，否则就没有意义——另一参与人会拖延使你达不到目的，或是一步一步地削弱你的威胁（腊肠策略）。这会使强迫威胁的实施比强迫许诺的实施更难。想让孩子努力学习的父母可以说："每个学期如果你可以拿到 B 或更好的成绩话，我会给你价值 500 美元的 CD 或电脑游戏。"小孩因此会更积极地达到每学期的目标。而以威胁来达到同一目的——"如果每学期你的成绩低于 B，我将会拿走你的电脑游戏"——就要求父母更主动更有警觉性，小孩会延迟把成绩单拿回家或把电脑游戏藏起来。

是奖励还是处罚与现实情况有关。如果小孩有权利永久拥有游戏的话，则拿走一部游戏就算是惩罚了；如果游戏是按学期发给小孩的，则下一学期继续发给他就是奖励。所以，你可以通过改变现状来将威胁转变为许诺，或将许诺变成威胁。你可以在进行策略性行动时通过这种改变获得好处。如果你想达到强迫，则制造这样一种场景，使当对方行动符合你的意愿时，你所做的是给予奖励，这时你使用的就是强迫许诺。举一个戏剧性的例子，强盗可以把威胁"如果你不把你的钱包给我的话，我就会杀死你"改成许诺"有一把刀已经架在你脖子上了，只要你把钱包给我，我就不杀你"。但是，如果你想达成遏制，则要营造另一种场景，使当对方行动违反了你的意愿时，你所做的是给予惩罚，此时你使用的就是遏制威胁。

10.6 建立置信度

我们已经强调了策略性行动置信度的重要性，在每个例子中也简略地说明了如何建立置信度。其方法通常是因事而异的，而且探索和发展这些方法是门学问很深的艺术。一些一般性原则可以帮你组织你的研究工作。

这里给出了建立置信度的两大方法：(1) 减少你未来行动的自由，那样你就没有选择，只能实施策略性行动中所确定的行动；(2) 改变你未来的支付，那样实施策略性行动会变成你的最佳行动。现在我们阐述如何具体使用这两大方法。

10.6.1 减少你行动的自由

1. 自动履行

假设你在第一阶段时做一些事使自己在第二阶段无法做出自由的选择，例如把它交给某个执行机制，该机制会在恰当情况下实施你的承诺、威胁或许诺，同时向对方证明你已经放弃选择的自由了，让他相信你不可能再改变想法了，这样你的策略性行动就是可信的。最有名的例子是**世界末日装置**（doomsday device）。它是一种引爆核弹的装置，只要敌人发动了核弹攻击战，它就会引爆并污染全球。这是 20 世纪 60 年代初期风行一时的电影《失败的防卫》（*Fail Safe*）和《奇爱博士》（*Dr. Strangelove*）的情节。幸运的是，这只是电影中的情节。但是在贸易政策中，如果别的国家试图对出口到你们国家的商品进行补贴，你们就可以启动自动报复程序，提高进口关税（反补贴税），这是很常见的。

2. 委托

执行机制不一定是机械式的。你可以将权力委托给其他人或组织，要他们按照特定程序或规则来做。事实上，这正是抵制关税的做法。美国政府有两个机构——美国商业部和国际贸易委员会，它们的运作程序是依照国家贸易法进行的。

被委托的代理人不能有自己的主见，以免无法达到策略性行动的目的。例如，某代理人被委托实施威胁惩罚，但代理人有暴力倾向，则他就有可能在对方如你所愿行动时，仍惩罚对方。如果对方预料到了这一点，则威胁就会无效，因为他知道无论如何都会受到惩罚。

委托机制不可能全然可信。即使是世界末日装置也可能不可信，因为对方会怀疑你设置了某个开关来控制装置以避免灾难的发生。委托也可以被取消。事实上，美国政府常常不考虑反补贴税而和其他国家达成某种协议，以避免贸易战。

3. 烧毁桥梁

很多入侵者，从古希腊的色诺芬（Xenophon）、英国的征服者威廉（William），到墨西哥的柯蒂斯（Cortes），都曾经故意切断自己军队撤退的路，破釜

沉舟，背水一战。有的是烧毁撤退的桥梁，有的是烧船，但这都是老话了。最近的一次应该是第二次世界大战中日本的自杀式战斗机只携带单程的油料去攻击美国的船只。这一原则在中国古代也曾出现过，孙子曰："兵士甚陷则不惧，无所往则固，深入则拘，不得已则斗。"夫差就说过"困兽之斗"的道理（"猛兽走到绝境，会奋不顾身反抗，真正的男子汉应该像这样！因为他们知道没有退路，只能奋战到底"[5]）。

赌注很高的博弈中，人们也会用类似的方法。虽然欧盟各国可以拥有自己的货币，而且有固定的汇率，但它们还是采用单一币制，这一不可逆的过程会让联盟成员更致力于联盟的发展。当然，放弃单一币制回到原本各国币制并非不可能，只是代价太大。如果联盟的效果真的很差，越来越多的国家可能会选择退出。所以，烧毁桥梁的置信度不是"全有或全无"，而是有程度差别的。

4. 切断通信

如果你发出一个信号向对方表明你的承诺，同时又切断和他的通信，则他就不能和你争论而让你收回承诺。切断通信的危险在于，如果双方同时这么做的话，则有可能造成双方承诺不一致，带来更大的互相伤害。另外，切断通信很难执行威胁，因为只有与外界保持联系，你才能知道对方的行动是否如你所愿，你是否需要实施你的威胁。当然在现在，要断绝一个人和外界的联系是十分困难的。

但是如果参与人是由团队组成的，则可以尝试这种方法的变形。比如劳工联盟，他们要开会决定工资。开一个会要有很多的计划——定会议室、联络会员等，需要几个礼拜的时间。如果工资没有达到劳工联盟的要求，联盟领袖就会号召罢工，那么就要再开一次会来决定新工资。这个讨价还价的过程让资方有了时间压力，因为他们知道劳工联盟几个礼拜才开一次会。所以切断一阵子的通信会带来一定的置信度，但绝不是绝对可信，联盟的通信只是延后了几个礼拜，而并非是断绝了。

10.6.2 改变你的支付

1. 声誉

通过实施威胁和实现许诺，你可以获得某种声誉。在与同一参与人的重复博弈中，这样的声誉是最为有用的。在与不同的参与人的不同博弈里，它也是有用的，如果他们中的每一个能够观察到你与其他人的博弈里的行动的话。有利于这样的声誉出现的情形与囚徒困境中达成合作的情形是一样的，并且有同样的理由。互动继续下去的可能性愈大，未来较之于现在就愈重要，参与人就愈可能牺牲当前的利益以获取未来收益。参与人就因此更加愿意去获取并维持声誉。

用专业术语来讲，这种机制将不同的博弈联系起来了，并且一个博弈里的行动的支付会因其他博弈中的后果之前景而改变。如果你在一个博弈里未能兑现你的威胁或者许诺的话，你的声誉将受到损害且你将在其他博弈里得到一个低的支付。因此，当你考虑这些博弈里的任何一个的时候，你就应该调整你在其中的支

付并且考虑到其对你在相关博弈里的支付造成的影响。

在持续的关系中好的声誉会为你带来利益。这就是为什么你经常光顾的汽车修理商基本上不会像路边摊一样漫天喊价，以次充好地欺骗你的原因。但如果竞争激烈，使他不得不以低价格揽客，而利润又微薄得几乎无利可图时，他又怎么从声誉中获取好处呢？所以他诚实帮你修车的前提是必须比同行收取的最低价格要高一点点。

同样的道理可以解释为什么当你远离家的时候，你会选择在知名连锁店用餐而不是在不知名的餐馆用餐。百货商店如果要推销新商品，它会用自己的声誉来许诺新商品的高品质。

一方或双方做出可信许诺的博弈可以给双方带来利益，参与人可以合作，建立并发展声誉。但是，如果相互影响只是在一个有限的时间内的话，则在博弈结尾会出现问题。

在 1993 年《奥斯陆协议》开始执行的中东和平进程的早期阶段，以色列将对加沙和约旦河西岸小片孤立地带的部分控制权移交给巴勒斯坦当局，而巴勒斯坦则承认以色列的存在并减少针对以色列的暴力活动，这种和平进程顺利地进行了一段时间。但是，当最后阶段到来时，双方下一步行动的置信度出现了问题，终于导致了和平进程于 1998 年中断。外界本可向双方提供富有吸引力的激励措施以使和平进程持续下去的，如美国和欧洲可向双方提供经济援助或扩大与它们的经贸往来规模。例如，美国曾以上述方式向埃及和以色列提供了大量的援助达成了 1978 年的《戴维营协议》(Camp David Accords)。但直到现在，仍没有任何类似措施被实施，因此，在写本书时，这一进程前景仍不容乐观。

2. 把博弈分成许多小块

有时候可以将单个的博弈分成许多连续的小博弈，好让声誉机制生效，就像建房工程中通常按工程进度分期付款一样。在中东和平进程过程中，以色列绝不会一次将整个西岸走廊交给巴勒斯坦以换回一个仅仅承认以色列的存在并停止恐怖活动的许诺。每一步的进展至少使进程部分地向前推进。但在博弈进行到尾声时会出现偶然性事件使博弈变得困难。

3. 团队合作

团队合作是将一个博弈嵌入另一个博弈以提高策略性行动置信度的另一种方法。它要求一组参与人互相监督。如果一个参与人没有实施威胁或兑现诺言，其他参与人就会要求惩罚他。如果不惩罚的话，那么连他们自己也要受到别人的惩罚，依此类推。所以在一个大的博弈中参与人的支付会被改变，从而使团队的置信度得到增强。

许多大学有学术荣誉规章，这样的规章就是一种置信度机制。考试的时候，学生不会受到监考，相反，学校要求学生如果看到作弊就要向学术委员会报告。委员会就会举行听证会，查证属实就进行惩罚。惩罚可能是休学一年或逐出校门。学生是不愿陷同学于不义的，为了坚定学生的决心，规定中加入了这样一条：如果有人发现某人作弊而不举报，也被视为违反规章。即使是这样，一般认为该规定还是不能达到完全的效果。普林斯顿去年的一个调查发现，只有三分之一的学生说他们愿意报告作弊，尤其是当他们知道是谁作弊的时候。

4. 非理性

你的威胁可能会缺乏置信度，因为其他参与人知道你是理性的，如果你实施了威胁，你付出的代价会很大。所以其他人认为，就算事到临头，你还是不会实施威胁的。你解决该问题的对策是宣称你是非理性的，使其他人相信你的支付和他们预想的是不一样的。当威胁的置信度受到怀疑时，表面的非理性就会转变成策略理性。同理，表面的非理性动机，像荣誉和面子，会使你的许诺可信，甚至使报复也变得可信。

其他参与人可能会洞悉这种理性的非理性（rational irrationality）。所以如果你试图通过宣称非理性来使你的威胁可信，他不会马上相信你。你必须获得非理性的声誉，例如，在相关的博弈中表现得非理性。当然你也可以使用第 9 章中谈到的策略，发出非理性的可信信号，从而将假的非理性剔除，达成一个均衡。

5. 契约

你也可以签下**契约**（contract），使你不实施威胁和兑现诺言的代价很大，因为违反契约的话是要付一大笔违约金的。如果该契约是明确的并且由法院或第三方权威机构执行，则支付的改变会让实施威胁和兑现诺言成为最佳行动，威胁和许诺就变得可信了。

至于许诺，契约的另一方是博弈的另一参与人。你兑现诺言对他是有利的，因而如果你不兑现诺言的话，他就会让你付出违约金。实施威胁的契约是有问题的。其他的参与人并不希望你实施威胁，也不愿意执行契约，除非他可以在相关的博弈中得到长期利益。因此，对于威胁，契约必须交由第三方执行。但是当引入第三方（执行方）时，契约仅仅能保证事到临头你会实施威胁；如果你不实施威胁，第三方就不能从中获益。所以契约就有再谈判的可能，以使第三方能从中获利。如果真到采取行动的时候，你可以对第三方说："看，我不想实施威胁。但因为违约金的缘故，我被迫这么做，而你不会得到任何好处。这里有一些钱，希望放我一马。"这样一来契约就不可信，从而威胁也不可信。第三方必须以长远利益为理由来执行契约，例如为保持其声誉，以保证契约没有再谈判的空间。

白纸黑字的契约比口头契约更有效力，但是有时候口头契约也能够构成承诺。当乔治·布什说："我承诺，以后不再课新税。"在 1988 年的总统大选中，美国民众把他的承诺视为契约，1990 年他征收了新税，民众在 1992 年大选中就没让他连任了。

6. 边缘政策

在美日贸易政策博弈中，我们发现威胁可能太大以至于都不可信了。但小而有效的威胁又不好找，这时可通过随机实施的方式来减小巨大的威胁，使其达到可信的水平。美国不可能跟日本说："你的市场不向美国开放的话，如果你遭到攻击，我不会前来保卫你。"但是它可以说："你不把市场向美国开放的话，这将会破坏我们两国的关系，并可能会带来风险。如果你面临入侵，我们的国会可能不会支持对你的军事协助。"如同之前说的，这种故意制造风险的方法称为边缘

政策。这是个微妙的策略，很难实施，看一个实际的例子最能够理解它的含义，第 14 章对古巴导弹危机的详细研究，对我们理解边缘政策会有所帮助。

我们已经介绍了一些使策略性行动可信的方法并讨论了它们如何发生作用。最后，我们想强调全部讨论所共有的一个特征。在实际中置信度并无有或无之分，只有强或弱之分。虽然没有什么理论支持（逆推法表明威胁要么会成功，要么就失败），但是实际应用时必须意识到在这两极中间仍存有一系列可能性。

10.7 反策略性行动

如果你的对手对你施加威胁或许诺，而他的威胁或许诺会对你产生不利影响，那么在他行动之前，你可以反策略性行动。这么做可以使他未来的策略性行动不那么有效。例如，破坏不可逆性和可行性。在本节中，我们将讨论一些可以达到这个目的的方法。同理，别人也可以用这些方法来达到他们自己的目的。

10.7.1 非理性

非理性对威胁或许诺的可能接受者一样有效。如果你被认为是如此地非理性，以至于不会屈服于任何威胁，并且愿意承受因对方实施威胁而受到的损失，那么在刚开始时他可能不会使用威胁，因为威胁的实施也可能会伤害到他自己。我们之前所说的说服对方相信你是非理性的很难，在这儿也可得到体现。

10.7.2 切断通信

如果你让对方不能将关于所做的威胁或许诺的信息传达给你，则他所做的一切就毫无意义。谢林用一个故事来证明这点：小孩啼哭的声音很大，以至于听不到父母的威胁。[6]那么父母做任何策略性行动都是没有意义的，因为通信已经被切断。

10.7.3 开辟逃走的路

如果对方能从烧毁桥梁阻止后退中获利，你也可以将火熄灭甚至修建新桥或新路让对方后退从而获利。这个方法古代人很早就知道了。孙子说："围师必阙。"当然，这么做的目的不是让敌人逃跑，相反是说明有一条路可以安全撤退，也就是让他知道除了死亡之外还有第二种选择，然后你再进攻。[7]

10.7.4 消除对方想要维持声誉的动机

如果有人威胁你说："看，我不想实施这个威胁，但我必须实施因为我想保持我的声誉。"你可以回应："我不乐于将你不处罚我的事公布于众，我只想在博弈中取得好的结果，我会保密，这会避免双方都受伤害的结果。你在别人面前的声誉也可以保持。"同理，如果你是买家，你和卖家讨价还价，他拒绝降低价格，他的理由是："如果我以这个价格卖给你，那我也必须以这个价格卖给别人。"你可以说你不会告诉其他人。这可能没用，对方可能怀疑你会告诉你的一些朋友，而你朋友又会告诉其他人。

10.7.5 腊肠策略

腊肠策略是削弱对方威胁的方式，就像腊肠每次被一片一片切去一样。你违背他人意愿（无论遏制或强迫）的程度非常小，以至于对方觉得他不值得为你的这种行为去实施对双方都不利的威胁。如果你的这种策略起了作用的话，你就可以继续一点一点地违背。

你在童年就很清楚这种做法。谢林[8]给出了这个过程的精彩描述：

> 我们确信，腊肠策略是由一个小孩创造的。告诉小孩不要下水，他会坐在岸边，把脚放在水里，这还不算下水。得到默许后，他就会站起来，将脚在水中探得更深一点。过了一会，他开始涉水，但不去很深的地方。可是他感觉这与刚才并没什么不同，于是他会向稍深的地方走去，并辩解说他只是在附近来回走罢了。不久后我们就叫他别游得太远，然后怀疑对他的告诫哪里去了。

腊肠策略对强迫极其有作用，因为可以利用时间优势。当你的父母让你打扫你的房间时，你可以推迟一个小时，说你要做作业，然后推迟半天，因为你要进行足球练习，然后是一晚，因为你不可能错过《辛普森一家》的电视节目，等等。

要对付腊肠策略，你的威胁处罚必须分等级制定，即必须有一定程度的处罚来对应一定程度的违反行为。也可通过逐步提高风险达到对付腊肠策略的目的，这也是边缘政策的另一种应用。

10.8 总　　结

为了制定后面的行动规则而采取的行动被称为策略性行动。真正的先行动必须是可观察的、不可逆的。如果先行动者试图获得想要的改变博弈均衡结果的成

效，那么先行动必须可信。承诺是无条件的先行动，只要先发优势存在，承诺就可以抓住先发优势。先行动通常包含着对一个策略的承诺，但这个策略并非原始版本博弈中的均衡策略。

有条件的先行动，如威胁和许诺，称作反应规则，或是遏制对方的行动以保持现状，或是强迫对方如你所愿地行动以改变现状。实施威胁可能会使双方都付出代价，但威胁成功的话就不用付出任何代价。如果威胁仅仅制造出逐渐升级的风险，就被视为边缘政策。对于许诺，只有许诺人得偿所愿时，许诺人才会付出代价。威胁可以任意大，虽然太大了就会不可信；许诺只需大到使其有效即可。如果威胁（许诺）中暗含的许诺（威胁）不可信，那么参与人必须做出明确可信的许诺和威胁，并采取使两者能相互结合的行动。

任何策略性行动都要建立置信度。有一些一般原则就是讲如何使策略性行动可信的，还有很多方法讲如何获得置信度。你可以通过减少你未来选择的自由，也可以通过改变你自己未来行动的支付来使策略性行动有效。具体包括建立声誉、团队合作、证明“表面”的非理性、烧桥断后路、签契约等。同样的方法适用于应对对方的策略性行动。

关键术语

边缘政策（brinkmanship）
承诺（commitment）
强迫（compellence）
契约（contract）
置信度（credibility）
遏制（deterrence）
世界末日装置（doomsday device）
不可逆的（irreversible）
可观察的（observable）
许诺（promise）
理性的非理性（rational irrationality）
反应规则（response rule）
声誉（reputation）
腊肠策略（salami tactics）
策略性行动（strategic moves）
威胁（threat）

习　　题

1. “有人说许诺的大小是有限制的，而原则上威胁可以是任意大的，只要它是可信的。”首先，简要解释一下为什么这段叙述是正确的。尽管这种叙述是正确的，但参与人可能发现使用过分严重的威胁对他们也未必有利。解释为什么这句话也是正确的。
2. 电影《曼哈顿谋杀谜案》（*Manhattan Murder Mystery*）中有这样的情景，伍迪・艾伦（Woody Allen）和黛安・基顿（Diane Keaton）正在麦迪逊花园看曲棍球比赛。很明显她不喜欢看，但是他告诉她：“记住我们的约定，你要在这儿陪我看完整场曲棍球比赛，下个星期我将会和你去看歌剧，陪你到底。”再后来，我们看到他们走出歌剧院，进入了冷清的林肯（Lincoln）中心广场，

广场里面还在放着音乐。基顿明显很生气："我们的约定呢？我陪你看完了曲棍球比赛，所以你也应该陪我看完歌剧。"艾伦回答："你知道我不能听太多的瓦格纳（Wagner）。在第一幕的结尾，我就已经很想波兰。"用你学到的策略性行动和置信度的理论知识来评论这里的策略选择。

3. 这里是父母和小孩之间的博弈。小孩可以选择表现好（G）或表现坏（B）；父母可以选择惩罚（P）和不惩罚（N）小孩。小孩从其坏行为中可得到价值为1的快乐，但受到价值为－2的惩罚。这样，表现好的小孩不会受到惩罚而得到支付0；而表现坏且受到惩罚的小孩会得到支付1－2＝－1，等等。父母则从小孩的坏行为中得到－2，从实施惩罚中得到－1。
   (a) 设定这个博弈为同时行动的博弈，找出均衡。
   (b) 然后，假设小孩会先选择G或B，而父母在观察到小孩的行为后会选择P或N。画出博弈树，找出子博弈完美均衡。
   (c) 假设在小孩行动前，父母能够采取某种策略性行动——例如，威胁如果你做坏事，我就惩罚你。父母拥有多少这种策略？写下这个博弈的支付表。找出所有纯策略纳什均衡。
   (d) (b) 与 (c) 的答案有什么不同，并解释其不同的理由。

4. 对于下面三个博弈中的每一个，回答这些问题：(1) 如果没有人使用策略性行动，如何达到均衡？(2) 一个参与人是否能通过使用一个策略性行动（承诺、威胁或许诺）或是联合使用这些策略性行动来提高他的支付？如果可以，两人中谁会使用什么样的策略性行动？

(a)

		列	
		左	右
行	上	0，0	2，1
	下	1，2	0，0

(b)

		列	
		左	右
行	上	4，3	3，4
	下	2，1	1，2

(c)

		列	
		左	右
行	上	4，1	2，2
	下	3，3	1，4

5. 修昔底德（Thucydides）所著的《伯罗奔尼撒战争史》中的策略博弈已经被圣路易斯大学的威廉·查尔依（William Charron）教授用博弈理论术语表示出来了。[9]雅典获得了爱琴海（the Aegean）周边的大的海岸城市，成为了一个大帝国，这些城市在保护希腊不受波斯人入侵中起到了主要的作用。斯巴达害怕雅典的实力，正在计划着对雅典的战争。如果斯巴达决定打的话，雅典就不得不考虑是保持还是放弃它的帝国。但是雅典也害怕，如果这些城市独立，它们可以选择加入斯巴达，成为一个联盟来反对雅典。这些城市这样做可以从斯巴达那里得到好处。所以有 3 个参与人——斯巴达、雅典、小城市，有 4 个结果，支付如下（支付 4 是最好的）：

结果	斯巴达	雅典	小城市
战争	2	2	2
雅典仍是帝国	1	4	1
小城市加入斯巴达	4	1	4
小城市保持独立	3	3	3

(a) 画出博弈树并找到逆推均衡，是否存在另外的更有利于所有参与人的结果。

(b) 哪种策略性行动或策略性行动的联合能获得更好的结果，讨论这种策略性行动的置信度。

6. 在经典电影《玛丽·波平斯》（*Mary Poppins*）中，小孩在和很多不同的保姆进行策略博弈。在他们的眼中，保姆本质上就是无情的，并且和保姆开玩笑又是很好玩的。也就是说，他们把自己看做一个博弈的参与人。在这一博弈中，保姆是先行动者，她们既可能表现得无情，又可能表现得和蔼；小孩后行动，他们可以选择是很乖的，也可以是很淘气的。保姆很乐意照顾乖的小孩，但仍然很严厉，所以她从（严厉，很乖）这个组合中得到最高支付 4，从组合（和蔼，淘气）中得到最低支付 1，从组合（和蔼，很乖）得到支付 3，从组合（严厉，淘气）得到支付 2。同理，小孩最希望的是一个和蔼的保姆，那样自己可以很顽皮。如果保姆和蔼的话，小孩能得到两种较好的结果（如果顽皮得到支付 4，如果很乖得到支付 3）；当保姆很严厉时，他们会得到较低的两种支付（顽皮得到支付 2，很乖得到支付 1）。

(a) 画出这个博弈的博弈树，找出没有策略性行动时的完美博弈均衡。

(b) 电影中，在玛丽·波平斯（Mary Poppins）来到之前，小孩在找保姆的广告中写道："如果你不责骂我们，我们也不会让你恨我们，我们不会把你的眼镜藏起来，不会把癞蛤蟆放在你的床上，不会把胡椒粉放在你的茶里。"使用（a）中的博弈树来说明这个陈述组成了一个许诺，如果这个许诺成立的话，会有什么样的结果？

(c) （b）中的许诺中暗含的威胁是什么？暗含的威胁自动可信吗？解释你的

答案。

(d) 孩子们怎么才能使在 (b) 中做出的许诺可信?

(e) (b) 中的许诺是强迫还是遏制? 可根据博弈的现状来解释你的答案，即如果不采取策略性行动会发生什么?

7. 重新确定习题 6 中博弈的支付，使得在广告中小孩的陈述是威胁而不是许诺是可能的。

(a) 重新画出习题 6 中的博弈树，填写双方的支付，使小孩的陈述变成完全意义上的威胁。

(b) 确定博弈中的现状，并决定使用哪一种威胁 (遏制还是强迫)。

(c) 解释为什么威胁不是自动可信的，给出支付结构。

(d) 解释为什么暗含的许诺是自动可信的。

(e) 解释为什么小孩开始时想使用威胁，并给出一个方法使其威胁可信。

8. 假设有两个国家 A 和 B，双方都有两种策略的选择："侵略"或"克制"。A 想要达到主宰世界的目的，所以侵略是其占优策略。B 想阻止 A 达到此目的，所以 A 侵略它就侵略，A 不侵略它就不侵略。[10] 支付表如下：

		A	
		克制	侵略
B	克制	4，3	1，4
	侵略	3，1	2，2

对于每一个参与人来说，4 是最好的支付，1 是最坏的支付。

(a) 找出同时行动纳什均衡。

(b) 下面考虑序贯行动博弈中的三种方式：(1) B 先行动，A 后行动；(2) A 先行动，B 后行动；(3) A 先行动，B 后行动，但 A 可以通过进一步的行动使自己后行动。为每一种方式画出博弈树，并找出子博弈完美均衡。

(c) 对于双方而言策略的关键在哪? (承诺、置信度等)

9. 考虑下面的博弈，在每个例子中，(1) 确定哪个参与人可以从做出策略性行动中获利；(2) 确定适合于这个目的策略性行动的本质；(3) 讨论在使这些策略性行动可信的过程中所产生的理论和实践上的困难；(4) 讨论这些困难是否能克服以及如何去克服。

(a) 欧盟中的其他国家 (法国、德国等) 希望英国加入单一货币和单一银行体系。

(b) 美国希望朝鲜停止向伊朗等国出口导弹及导弹技术，希望中国参与进来和美国一起来达到此目的。

(c) 美国汽车工人希望美国汽车生产者不要在墨西哥建厂，并希望政府对外国汽车的进口进行限制。

(d) 你们大学的学生想要阻止学校提高学费。

(e) 许多的参与人还有局外人都希望北爱尔兰和中东能够长期和平。

10. 简要介绍一场你自己曾经参加过的博弈，其中应包括一系列的策略性行动，

例如承诺、威胁等，特别要注意置信度的重要性。如有可能，描述一下这场博弈，并解释为什么会有这种结果。参与人是否在好的策略性思维指导下做出了合理的选择？

【注释】

[1] 我们在第 8 章和第 13 章都可以看到此博弈存在第三个均衡，也就是混合策略均衡。

[2] 你可能不同意这些结果的排名，也许既不适合你，也不适合你的老师。但希望你能暂时接受它，它的主要目的是以一种简单的方式来传达关于承诺的一般观点。接下来的例子同样适用这一声明。

[3] 这里再次地请你们接受这样的为了表达思想的支付结构，你可以通过改变支付表来看一看角色的转换以及策略性行动的有效性受到了怎样的影响。

[4] 在第 11 章中，我们将会更仔细地研究囚徒困境中企图达到合作结果的重复和持续的关系的重要性。

[5] Sun Tzu, The Art of War, trans. Samuel B. Griffith, Oxford: Oxford University Press, 1963, p. 11.

[6] Thomas C. Scheling, *The strategy of Conflict*. Oxford University Press, 1960, p. 146.

[7] Sun Tzu, *The Art of War*, pp. 109 - 110.

[8] Thomas C. Schelling, *Arms and Influence*, New Haven: Yale University Press, 1966, pp. 66 - 67.

[9] William C. Charron, "Greeks and Games: Forerunners of Modern Game Theory," *Forum for Social Economics*, vol. 29, no. 2, spring 2000, pp. 1 - 32.

[10] 感谢加州大学洛杉矶分校（UCLA）的政治科学教授 Thomas Schwartz 给出了该习题的思路。

第 11 章　囚徒困境和重复博弈

本章将继续分析更为广泛意义上的囚徒困境。大多数人或许都听过囚徒困境，哪怕对博弈论毫无所知者也对此故事略知一二。对于博弈论及其对参与人行动的预测来说，囚徒困境可谓经典的例子。在囚徒困境中，每个参与人都有占优策略，但均衡时（参与人都采用占优策略）的支付比参与人都选用劣策略时的支付要低。这个均衡的矛盾的性质引出了关于博弈互动更为复杂的问题，只有更全面的分析才能解答。本章旨在做这方面的介绍。

4.3 节已经讨论过囚徒困境。那时我们就发现均衡的结果对于参与人来说都是不好的，"囚徒"可以找到比均衡更好的结果，但却很难放弃均衡策略来实现这个结果。本章的焦点是研究达到更好结果的可能性，也就是考虑能否以及如何克服"囚徒"追逐私利的背叛动机，让他们达到并维持互利的合作结果。我们首先回顾标准的囚徒困境博弈，然后研究三种解决方法。第一种也是最重要的解决方法就是单次博弈的重复，另外两种可能的解决办法就是依赖于惩罚（或奖励）规则和领导角色。当我们考虑每一种解法时，背叛的成本和合作的收益就显得非常重要了。

最后，本章会讨论囚徒困境的实验证据，以及现实世界中囚徒困境的例子。实验的设计者会让实验的参与人处于各种类型的囚徒困境中。其所揭示的行为，一些是让人费解的，而另一些则是可以预测的；借助计算机模拟的实验可以得到另外一些有趣的结果。本章最后会提供现实世界中囚徒困境的例子，让读者认识到囚徒困境出现的场合是非常丰富的，并且至少有一个例子展示了参与人是如何

创造出自己解决困境的方案的。

11.1　基本博弈（复习）

在讨论避免囚徒困境中的“坏”的结果的方法以前，我们先简单地回顾一下这种博弈的基本情况。我们曾在第4章中讨论过丈夫和妻子涉嫌谋杀的例子，他们被警方隔离审讯并且可以选择承认犯罪或否认犯罪。我们将夫妻双方面临的支付矩阵表4—4复制到了表11—1，表中的数字代表了被判入狱的时间（年），数字越小对博弈双方就越有利。

双方都有其占优策略，也就是无论对方如何选择，自己都选择“认罪”。均衡结果是双方都选择“认罪”且均坐10年牢，但如果他们都选择“抵赖”，结果会更好一些，他们各自只要坐3年牢。

在任何囚徒困境中，都会有合作策略和欺骗或背叛策略。在表11—1中，“抵赖”是合作策略，双方都能获得最佳的结果；而“认罪”是欺骗或背叛策略，当参与人不相互合作时，他们会选择“认罪”，以牺牲对方利益换取自己的利益。所以囚徒困境中的参与人又可以根据他们策略选择的不同被分为背叛者和合作者，我们将在关于解决囚徒困境的可能途径的讨论中一直使用这个分类称呼。

表11—1　　　　**标准囚徒困境的支付**

		妻子	
		认罪（背叛）	抵赖（合作）
丈夫	认罪（背叛）	10，10	1，25
	抵赖（合作）	25，1	3，3

有一点值得强调，虽然我们在说合作策略，但是囚徒困境是我们在第2章中就解释过的非合作博弈，即，参与人策略的选择和实施是独立的。如果两个参与人可以一起探讨、选择和执行他们的策略——比如，两个参与人被关在同一个房间，并且可以就是否承认犯罪一起作答，这样就很容易达到他们想要的均衡结果。囚徒困境能否、何时以及如何解决这一问题，关键就难在要通过非合作（独立的）行为达到合作（共同较优的）结果。

11.2　解法Ⅰ：重复

在囚徒困境中所有可以维持合作的机制里，最著名且最自然的方式就是**重复博弈**（repeated play of the game）。参与人之间重复和持续的关系是该博弈的特

征。在囚徒困境中，参与人会担心一有背叛就会导致未来合作的崩溃。如果未来合作产生的支付比短期背叛的支付要大，则从个人长期利益着想，参与人自动不会选择背叛，同时这也不需要动用第三方来施加惩罚或压力。

我们这里来讨论两家餐馆泽维尔的达帕斯西班牙餐馆和伊冯娜的炭火西餐馆所面临的餐费定价困境。这个博弈曾在第 5 章介绍过，但这里从讨论的目的出发，我们简化这个博弈，假设只有两种价格可供选择：共同最优（共谋）的价格 26 美元和纳什均衡的价格 20 美元，每家餐馆的支付（以百美元/月为单位计量的利润）可以通过使用 5.1.1 节的数量（需求）方程计算得到，结果如表 11—2 所示。与任何囚徒困境一样，虽然这两家餐馆都更偏好互相合作和将餐费定在高价 26 美元的结果，但它们都会选择占优策略“背叛”且将餐费定在 20 美元。

表 11—2　　定价的囚徒困境（百美元/月）

		伊冯娜的炭火西餐馆	
		20（背叛）	26（合作）
泽维尔的达帕斯西班牙餐馆	20（背叛）	288，288	360，216
	26（合作）	216，360	324，324

分析一开始，让我们设想两家餐馆最初是合作模式，都以高价 26 美元经营。如果某一餐馆，如泽维尔，背离这个价格策略，那么它可以将支付从每个月 324 提高到每个月 360（从 32 400 美元提高到 36 000 美元）。但是从此合作关系就被破坏了，泽维尔的对手伊冯娜也没有理由再继续合作。一旦合作关系破裂（假设是永远破裂了），泽维尔的支付从原来它没有背叛价格合作时的每月 324 降到了现在的每月 288，虽然在背叛的那个月得到了 36（3 600 美元）的收益增量，但从此就得为破坏合作付出每个月 36（3 600 美元）的损失。即使它们之间的关系只存在三个月，选择背叛也不是泽维尔的最优选择。伊冯娜也可以做类似的分析。因此，如果两个餐馆竞争超过三个月，我们看到的就是合作行为的高价，而不是单次博弈预测的背叛行为的低价。

11.2.1　有限次重复

但囚徒困境的解决实际上没那么简单，如果它们只经营三个月呢？则餐馆会分析并且选择它们三个月中的最优策略。每家餐馆都会用逆推法来确定每个月的定价。从第三个月的分析开始，它们了解到此时已经没有进一步的关系值得考虑了，每家餐馆都会选择占优策略背叛。因此，它们在第二个月也会认为双方没有更进一步的关系，双方知道在第三个月会互相背叛，所以在第二个月也会背叛。同样的道理应用到第一个月，知道第二个月和第三个月会互相背叛，它们会认为在第一个月没有合作的价值，所以会从第一个月开始就背叛，困境因此又出现了。

这一结果很普遍。在囚徒困境中，只要参与人间的关系维持一段有限的时

间，占优策略背叛就会在最后一期被采用。一旦参与人到达博弈最后阶段，就没有继续合作的价值，所以就会产生背叛，而逆推法预测背叛会从博弈最后一期蔓延到第一期。然而，实际上有限次重复的囚徒困境却有很多合作的局面出现。

11.2.2 无限次重复

对有限次重复囚徒困境的分析证明了即使博弈重复进行也不能解决困境，但是如果双方的关系没有一定的期限又会怎么样呢？如果两家餐馆彼此永远竞争下去呢？我们在分析中必须加进这些相互联系的因素，我们会发现参与人的动机也会随之改变。

在任何重复博弈中，参与人行动的有序性意味着他们在本回合中采取的策略取决于博弈前一回合中的行动，这样的策略被称为**条件策略**（contingent strategy），在重复博弈的理论中有几个特别的例子会常常被用到。大部分的条件策略是**触发策略**（trigger strategy），使用触发策略的参与人只要对手合作的话，会跟对手合作，但只要有任何背叛就会触发一段**惩罚期**（punishment），他会采取不合作的态度来回应。两个最有名的触发策略是冷酷策略和以牙还牙策略。**冷酷策略**（grim strategy）是指直到对手背叛你之前，一直和他合作；而一旦他背叛你，你就在剩下的博弈中（通过选择背叛策略）惩罚他。[1] **以牙还牙策略**（Tit-for-Tat，TFT）不像冷酷策略般不原谅对手，它因不需要永久性地惩罚对手来解决囚徒困境而得名。以牙还牙策略是指对手合作你就在近期也合作，对手背叛你就背叛，惩罚期限依据对手是否继续背叛而决定，并且在他选择重新合作后，你也选择重新合作。

思考一下，当某家餐馆使用以牙还牙策略时，重复博弈是如何进行的。我们已经知道当泽维尔的达帕斯西班牙餐馆背叛合作一个月，它会额外获得 36 的支付（360 而不是 324），但如果泽维尔的对手采用以牙还牙策略，则伊冯娜的炭火西餐馆会在下一个月惩罚泽维尔来进行报复。此时泽维尔有两个选择，它可以继续以 20 美元的价格背叛合作，而伊冯娜会根据以牙还牙策略来惩罚它，这样的话，泽维尔会在接下来的每个月失去 36 的支付（288 而不是 324）。这种选择看起来是代价很大的，但泽维尔也可以回到合作策略来。所以在背叛后的第一个月，泽维尔回到合谋的价格 26 美元上来，它只受到了伊冯娜一个月的惩罚，并且在受惩罚中支付损失 108（它得到 216 而不是不背叛时的 324）。在泽维尔背叛后的第二个月，博弈双方都回到合作的价格上来并得到每月 324 的支付。这个一次的背叛行为产生了额外 36 的支付，但在受惩罚期间也额外付出了 108 的损失，这对于泽维尔来说代价也是很大的。

这里有一个重点要认识到，泽维尔在第一个月背叛得到额外的支付 36，但未来会遭受损失，所以这一得一失的相对重要性取决于现在与未来的相对重要性。因为支付是以百美元计算的，所以我们可以做客观比较。一般而言，现在赚的钱（利润）比以后赚的钱要更有价值，因为你可以将钱拿去投资并获取利息，所以泽维尔应该依照投资的收益率（包括资本利得、股息或利息，依投资类型而

定）来计算背叛是否划算。我们用符号 r 来代表收益率，因此投资 1 美元将带来 r 美元的利息、股息或资本利得，或者投资 100 美元产生 $100r$ 的利息等，那么收益率又可以记为 $100r\%$。

我们可以计算背叛是否合乎泽维尔的利益，因为餐馆的支付是以百美元计算的，而不是像前面的章节那样以等级计算（例如第 3 章和第 6 章的街心花园博弈）。这就意味着不同单元格中的支付是可以直接比较的，4（美元）的支付价值是 2（美元）的两倍，但在任何二阶博弈中，当四个不同的结果用 1（最差）到 4（最好）来衡量等级时，支付 4 就不一定是支付 2 的两倍。只要支付是用可以衡量的单位表示的，就可以计算囚徒困境中的背叛是否划算。

（1）对手采取以牙还牙策略时，背叛一次是否划算？

在泽维尔的对手使用以牙还牙策略时，它的一个选择是背叛一次然后再回到合作模式，这一策略让泽维尔在第一个月（背叛的那个月）额外得到 36 的利润，但在第二个月失去 108 的利润，第三个月恢复合作。那么，仅背叛一个月是否划算呢？

我们不能直接拿第一个月的 36 和第二个月的 108 相比，因为在计算中必须考虑资金的时间价值。也就是说，我们需要找到一种方法决定第二个月损失的 108 在第一个月值多少钱，然后拿来跟 36 比较，看看背叛一次是否值得。我们找的是 108 的**现值**（present value，PV），或者说是下个月赚的 108 在这个月（现在）等价的值为多少。我们需要计算这个月要赚多少，加上利息，才能在下个月等于 108，我们把这个数记为 PV，也就是 108 的现值。

假定（每月的）收益率是 r，即这个月得到数额为 PV 的钱，然后拿去投资到下个月，可以收回的资金为 $PV+rPV$，其中第一项是收回的本金，第二项是收益（利息、股息或资本利得）。当收回资金的总数正好为 108 时，PV 就等于 108 的现值。令 $PV+rPV=108$，解出 PV 得：

$$PV=\frac{108}{1+r}$$

只要给定 r，我们就会知道今天的多少钱在下个月值 108。

从泽维尔的角度来考虑，问题在于这个月得到的 36 能否弥补下个月的损失 108。这个问题的答案依赖于现值 PV。泽维尔必须把 36 和 108 的现值做比较，只有当 $36>108/(1+r)$ 时，也就是 $36(1+r)>108$ 或 $r>2$ 时，背叛一次才更划算。所以面对对手的以牙还牙策略时，泽维尔只有在每月收益率超过 200%时，才会选择背叛一次。显然，这是不可能的，比如年借款利率很少超过 12%，这转换为月利率就不足 1%（按年而不是按月记复利），远远低于刚才计算的 200%。所以在其对手执行以牙还牙策略时，泽维尔还是继续合作比较好，而不是选择背叛一次。

（2）当对手采用以牙还牙策略时，永久背叛是否划算？

如果一旦选择背叛就一直背叛下去，情况又是怎么样的呢？这是泽维尔的第二个策略选择。它会在第一个月得到 360，但如果对手一直执行以牙还牙策略的话，在接下来的每一个月中它会损失 36。要决定这个策略选择是否合乎泽维尔

的最佳利益，必须把未来损失的利润转换为现值，但这一次，未来的竞争导致的损失延伸到了**无限期**（infinite horizon）。

我们必须计算出所有这些未来损失的支付 36 的现值，然后加起来，再跟背叛当月的额外支付 36 去比较。泽维尔在选择背叛后继续选择背叛，它受到惩罚的第一个月将损失 36，其现值是 $36/(1+r)$，这与上述关于 108 现值的计算 $108/(1+r)$是一致的。再过一个月，也就是两个月后损失的 36，等于这笔损失的现值加上两个月的**复利**（compound interest）之和。如果今天投资 PV，那么一个月后这笔投资将变为本金加上收益 rPV，即总共为 $PV+rPV$；将这笔资金在第二个月继续投资，那么在第二个月末，投资者将得到第二个月开始时的（$PV+rPV$）加上投资这笔钱的收益 $r(PV+rPV)$。因此，两个月后损失的 36 的现值等于下面这个方程的解：$PV+rPV+r(PV+rPV)=36$，解出 PV 得 $PV(1+r)^2=36$，或者 $PV=36/(1+r)^2$。依此类推，在继续背叛的第三个月损失的 36 的现值是 $36/(1+r)^3$，第四个月损失的 36 的现值为 $36/(1+r)^4$。实际上，在继续背叛的第 n 个月损失的 36 的现值为 $36/(1+r)^n$。泽维尔会有无限个月的 36 的损失，每一个月的损失的现值都会越来越小。

更精确地说，泽维尔的损失是 $36/(1+r)^n$ 从 $n=1$ 到 $n=\infty$的加总（其中 n 代表开始背叛后继续选择背叛的第几个月），从数学上来说，这是无穷级数求和[2]：

$$\frac{36}{(1+r)^1}+\frac{36}{(1+r)^2}+\frac{36}{(1+r)^3}+\frac{36}{(1+r)^4}+\cdots$$

因为 r 是收益率，它是个正数，所以比率 $1/(1+r)$会比 1 小；这个比率通常称为**贴现因子**（discount factor），记做希腊字母 δ。当 $\delta=1/(1+r)<1$ 时，关于无穷级数的数学规律告诉我们这个和会收敛于某一个特定值，在这里就是 $36/r$。

现在可以确定泽维尔是否要选择永久背叛了。把 36 的额外支付和所有 36 的损失的现值（即 $36/r$）比较，只要 $36>36/r$，或$r>1$，它就选择永远背叛，即，只有月收益率超过 100%时，永远背叛才是更有利的。但这种情况是不可能的。因此，当双方都采用以牙还牙策略时，泽维尔最好不要背叛合作，双方都采取高价的合作结果才是纳什均衡。所以以牙还牙策略的采用解决了两家餐馆间的囚徒困境。

记住以牙还牙策略只是触发策略中的一种，它是比较“仁慈”的策略，所以如果以牙还牙策略能够解决两家餐馆间的囚徒困境，则其他更严厉的策略更应该可以，例如冷酷策略可以用来维持无限次重复博弈的合作关系。

11.2.3 未知期限的博弈

除了考虑有限期跟无限期的博弈外，我们还可以采用更加成熟的方法来处理未知期限的博弈。在某些重复博弈中，参与人可能不确定他们之间的互动会持续

多久。他们也许会认为博弈再持续一个时期是一个概率的问题，比如餐馆可能会认为只要它们的顾客将客饭定价菜单作为外出就餐的选择，重复博弈就会持续；而如果照菜单点菜以某种概率取代了客饭定价菜单，那么博弈的性质就会改变了。

我们已经知道，下一个月的损失的现值是 $\delta=1/(1+r)$ 乘以该损失，如果考虑到下一个月博弈会继续的概率 p(比 1 小)，则下一个月的损失的现值是 p 乘以 δ 再乘以损失的利润。对于泽维尔的达帕斯西班牙餐馆来说，这意味着当确定博弈会继续时，由于继续背叛而损失的 36 的现值是 $36\times\delta$，即 $36/(1+r)$，而当博弈会继续的概率是 p 时，则现值只有 $36\times p\times\delta$。因为 $p<1$，所以将博弈下一个月结束的概率纳入考虑范围意味着损失 36 的现值比博弈之中确定会继续时（此时 p 被假设等于 1）更小。

当我们将 p 纳入考虑之中时，未来支付的有效贴现因子是 $p\times\delta$，而不是 δ。我们将 R 称为**有效收益率**（effective rate of return），R 满足$1/(1+R)=p\times\delta$，R 跟 p 和 δ 有以下关系[3]：

$$1/(1+R)=p\delta$$
$$1=p\delta(1+R)$$
$$R=\frac{1-p\delta}{p\delta}$$

若实际收益率为 5%($r=0.05$，$\delta=1/1.05=0.95$)，博弈下一个月会继续的概率是 50%($p=0.5$)，则 $R=[1-(0.5)(0.95)]/(0.5)(0.95)=1.1$ 或 110%。

如果我们把收益率看成有效收益率，而不是实际收益率，破坏合作（鼓励背叛）所需的临界收益率就会显得更接近实际。可以想象，如果博弈在最近将会结束的概率很高，则永久背叛或一次背叛就可能对参与人更有利。当面对采取以牙还牙策略的对手时，考虑泽维尔决定是否永久背叛。我们之前的计算结果显示仅当 r 超过 1 或者 100%时，永久背叛才是有利的。和我们上一段的假设一样，如果泽维尔的实际收益率是 5%，而博弈在下个月会继续的概率是 50%，则有效收益率是 110%，这就超过了继续背叛的临界。如果博弈在下一个阶段结束的概率足够大——也就是 p 值足够小，基于以牙还牙策略上的合作行为就可能会瓦解。

11.2.4 一般理论

我们可以简单归纳一下当对手采取以牙还牙策略时确定是否值得背叛的条件，这样你就可以将其应用到你所遇到的囚徒困境。首先，我们以一般符号来表示支付（以可衡量的单位表示）。如表 11—3 所示，该支付表必须符合囚徒困境的标准结构，也就是 $H>C>D>L$，其中 C 是合作结果，D 是双方都背叛的结果，在一方合作而另一方背叛时，H 是背叛者的高收益支付，L 是合作者的低收益支付。

表 11—3　　　　囚徒困境的一般形式

		列	
		背叛	合作
行	背叛	D，D	H，L
	合作	L，H	C，C

在标准的囚徒困境中，参与人背叛一次获得（$H-C$），后来又回到合作而被惩罚一次损失（$C-L$），永久背叛时每期会损失（$C-D$）。更加一般化，我们把 p 纳入考虑之中，也就是下一期博弈以一定的概率 p（$p<1$）继续，所以我们将支付用有效收益率 R 贴现。如果 $p=1$，即博弈确定会继续，那么 $R=r$，就是我们在前面计算中用到的简单收益率。用 R 替代 r，前面得到的各结论就更加一般化了。

我们知道，当对手采用以牙还牙策略时，如果背叛一次得到的收益（$H-C$）超过了被惩罚一次损失的现值［即（$C-L$）的现值］，则参与人就会选择背叛。一般地，参与人背叛一次采取以牙还牙策略的对手，仅当$(H-C)>(C-L)/(1+R)$ 或$(1+R)(H-C)>C-L$，或者

$$R>\frac{C-L}{H-C}-1$$

同理，当对手采取以牙还牙策略时，如果永久背叛得到的一次收益超过了未来各期损失现值的无穷级数和［各期损失是（$C-D$）］，则参与人就会选择永久背叛，即$(H-C)>(C-D)/R$ 或

$$R>\frac{C-D}{H-C}$$

这些结果让我们知道，参与人在决定是否背叛时，有三个因素很重要，分别是背叛马上得到的收益（$H-C$），因惩罚而遭到的损失［每个惩罚期损失为（$C-L$）或（$C-D$）］，以及有效收益率的值（R，衡量现在和未来的相对重要程度）。那么，当这些变量满足什么条件时，参与人才有动机去背叛呢？

第一，假设由背叛导致的收益值和损失值保持不变，R 的大小不同就决定了参与人是否会选择背叛。当 R 很大时，背叛更有可能发生。大的 R 值与小的 p 值和小的 δ 值（大的 r 值）相关。在博弈继续的概率很小或贴现因子很小（或利率很高）的情况下，背叛则更有可能发生。另一种解释的方法是，当未来相对现在而言不那么重要时，或者当未来很渺茫时，背叛更有可能；也就是当参与人没有耐心或博弈有望很快就结束时，参与人更有可能背叛。

第二，当有效收益率和背叛的单次收益一定时，各期因惩罚而导致的损失则决定了背叛是否划算。当（$C-L$）或（$C-D$）越小时，参与人越有可能背叛。在这种情况下，当惩罚不是那么严厉时，背叛越有可能发生。[4]

第三，假定有效收益率和各期由于惩罚而导致的损失保持不变，则当背叛的收益（$H-C$）很大时，参与人更有可能背叛，当背叛带给参与人大而且快的收益时，更是如此。

这里的讨论也突出了发现欺骗行为的重要性。要决定是否继续合作，就必须知道欺骗行为能维持多久而不被发现，也必须知道它在何种准确程度上被发现，还必须知道恢复到合作前惩罚要维持多久。如果背叛行为可被准确而迅速地发现，则它的收益就不会持久，当然也更确定要为此付出代价。所以触发策略能否成功解决重复博弈的囚徒困境依赖于背叛行为多快多准地被发现，这也是以牙还牙策略不是很好的原因之一，因为一个在执行和观察背叛行为上的小的失误将会导致参与人陷入连环惩罚而不可自拔，直到相反的小失误发生为止。

使用这些理论作为指导，你可以预见什么时候对手间出现合作的行为，以及什么时候出现背叛行为。如果市场不景气，整个行业处于崩溃的边缘，那么厂商们就觉得没有未来，则竞争可能比平时更残酷（基本上看不到合作行为）。即使市场暂时还景气，但不会持久，厂商可能想着趁现在可以快点捞一笔，因而合作行为可能瓦解。同理，如果某个行业因一时流行而兴起，但当流行热潮退却时就会衰退，则也可能不会有合作行为。如某一海滩度假地变得很热门，但那里所有的酒店都认为这不会长久，则它们不会形成价格上的同盟。另一方面，如果流行产品的交替是发生在类似的有长期联系的厂商之间，则合作模式有可能存在。例如，即使所有的小孩某一年都喜欢抱抱熊，而另一年都喜欢动力骑兵之类的玩具，价格的合作仍有可能发生，只要这些厂商都生产这两类玩具的话。

第 12 章将会进一步介绍多个参与人参与的囚徒困境，并探讨何时可以及如何克服这些困境以达到对于参与人来说更好的结局。

11.3 解法Ⅱ：惩罚与奖励

虽然重复是解决囚徒困境的主要手段，但也有其他方法可以达到这个目的。一个最简单的解决单次博弈囚徒困境的方法就是施加直接的**惩罚**（penalty）于背叛的参与人。将惩罚的成本纳入到考虑之中后，参与人的支付就会发生改变，囚徒困境也就迎刃而解了。[5]

考虑 11.1 节中夫妻的囚徒困境。如果只有一位参与人背叛，博弈的结果就会让背叛者坐 1 年的牢，合作者坐 25 年的牢。虽然背叛者早些出狱，但会发现合作者的朋友已经在外面等着揍他。这身体的伤害可能等价于另外 20 年的徒刑。所以如果参与人考虑受伤的可能，则原博弈的支付结构就会改变。

考虑了身体惩罚的新博弈如表 11—4 所示，当一方“认罪”而另一方“抵赖”时，加 20 年额外的徒刑于认罪的一方，博弈就完全改观了。

表 11—4　　考虑对唯一背叛者惩罚的囚徒困境

		妻子	
		认罪	抵赖
丈夫	认罪	10，10	21，25
	抵赖	25，21	3，3

在表 11—4 中没有发现占优策略，由逐个单元格检查法得到两个纯策略纳什均衡，其中一个是（认罪，认罪），另一个是（抵赖，抵赖）。现在每个参与人都发现如果对方合作的话，最符合自己利益的策略就是跟着合作。这个博弈已经从囚徒困境变成了第 4 章的安全博弈，新博弈的求解即需要从两个均衡中选出一个。当然，从两个参与人的利益来看，其中合作的均衡比互相背叛的均衡要好，所以当预期收敛于此的时候，合作很容易成为聚点均衡。

注意，在上面的情形下，背叛者仅当其对手没有背叛时才被惩罚。然而，我们可以制定更为严厉的惩罚措施，例如，任何认罪行为都会受到惩罚。这一惩罚措施必须交由权威的第三方而不是某一参与人的朋友来执行，因为当该参与人也背叛时，该参与人的朋友会偏袒该参与人而倾向于惩罚另一参与人。如果两个囚犯都是某帮派分子（如黑手党成员），而该帮派对不向警察认罪有公正的规定，对背叛的惩罚也是极其残酷的，那么博弈的支付情况会再次变为如表 11—5 所示的结构。

表 11—5　　对任何认罪都惩罚的囚徒困境

		妻子	
		认罪	抵赖
丈夫	认罪	30，30	21，25
	抵赖	25，21	3，3

现在等价的额外 20 年徒刑加在所有“认罪”策略上（比较表 11—5 和表 11—1）。在新的博弈中，每个参与人都有占优策略，不同的是支付结构的变化使得“抵赖”成为他们的占优策略，而（抵赖，抵赖）成为唯一的纯策略纳什均衡。由第三方执行的严厉惩罚让参与人不再选择背叛，而使合作成为博弈的新均衡。

在规模更大的囚徒困境中，使用惩罚会有困难，特别是有很多参与人或有不确定情况发生时，惩罚措施会很难实施，认定是否存在背叛行为也变得很困难，因为有可能只是运气不好或行动错误。另外，如果真的有人背叛，通常很难从众人中辨别出背叛者的身份。而且如果是单次博弈，未来就没有机会制定更严厉的惩罚措施或施加惩罚于后来被识别出来的背叛者。所以相比于两人博弈，使用惩罚在规模较大的单次博弈中很难达到预期目的。第 12 章将进一步讨论具有众多参与人的囚徒困境。

虽然惩罚计划有可能解决囚徒困境，但会产生有趣的问题。如某些困境均衡结果对参与人不利，但对剩下的社会大众或其中的某一部分却是有利的。如果是这样的话，社会或政治压力可能会使参与人跳出囚徒困境的能力最小化。当我们用第三方的惩罚来解决囚徒困境时，例如帮派老大的制裁力量，他规定参与人必须否认犯罪，这时候社会有自己的策略来降低惩罚机制的有效性，联邦证人保护计划就是为达成此目的而建立的。美国政府排除了惩罚的威胁，以保护认罪的呈堂证供。

其他的囚徒困境中也可以看到类似的情况。比如两家餐馆的价格竞争博弈，均衡时两家餐馆都定低价 20 美元，即使在它们都定高价 26 美元时双方都会更好

一些。虽然两家餐馆都想打破这个"不利"的均衡——正如使用触发策略可以实现的一样，但它们的顾客却更偏好单次博弈中低价的均衡；所以顾客们会有动力来破坏餐馆用以解决困境的机制。比如，因为一些面临囚徒困境式定价博弈的厂商可以通过"友好竞争"、"统一价格"等运动来解决困境，消费者们可以施加压力为禁止这些运动立法。我们将在11.6.4节中分析这种"统一价格"运动的效果。

囚徒困境可以借助惩罚背叛者来解决，同时也可以通过奖励合作者来解决。因为这种解决办法实际上实施起来比较困难，所以我们只简单提一下。

最重要的问题是由谁提供奖励。如果是由第三方，那么在参与人合作的情况下这个充当第三方的个人或集团一定是有利可图的，这样提供奖励才是值得的。一个不常见的例子就是美国通过承诺向以色列和埃及提供援助，居间斡旋使双方达成了《戴维营协议》。

如果奖励由参与人自己提供，要诀是该奖励应是有条件的（对手合作的话才提供奖励）和可信的（对手合作的话一定提供奖励）。要达到这些标准需要一些特别的安排，例如，提供奖金的参与人应该预先把钱存在银行，由可敬和中立的第三方保管，如果对手合作的话就把钱给他，反之则把钱还给提供奖金的人。本章习题11会让你知道这一方法是如何起作用的，虽然这道题是人为设计的。

11.4 解法Ⅲ：领导

囚徒困境最后一种解决办法是让某参与人在博弈互动中充当领导者的角色。在大部分囚徒困境的例子中，博弈都被假定是对称的，也就是参与人在背叛（合作）时，他们失去（得到）的支付都相等。然而，在真实博弈中，有些参与人的支付可能很大（领导者），有些可能很小。如果支付相差足够大，则背叛时，大部分的伤害会落在有较大支付者的身上，这使得他即使知道对手会背叛，也会选择合作。例如，沙特阿拉伯在OPEC（石油输出国组织）中，扮演"摇摆者"的角色已经很多年，如果石油生产量小的国家（如利比亚）扩张输出的话，它就会减少输出，以保持高的油价。

就如OPEC的例子一样，国家之间的博弈较公司或个人之间的博弈更常出现**领导**（leadership），所以我们用来介绍领导解决囚徒困境的例子就是国与国之间博弈。设想多米尼加（Dorminica）和索婆里亚（Soporia）两国的人口受到一种名为急性发作睡眠（SANE）的疾病威胁，每2 000人中有1人会感染这种疾病，即0.05%的人口感染，并且这种疾病可以让感染者陷入为期一年的深度睡眠。[6]这种疾病没有后遗症，但是一个工人由于生病一年不能工作的损失是32 000美元。每个国家有1亿名工人，所以预计患病的人数是50 000（0.000 5×100 000 000），预计的损失是16亿美元（50 000×32 000），全球的预计损失——多米尼加和索婆里亚的损失之和——是32亿美元。

科学家确信投入20亿美元到一个研究项目，可以迅速地研制出一种100%有效的疫苗。比起全人类因疾病而遭受的损失而言，这个研究项目是值得开展

的。然而，每个国家的政府必须独自决定是否出钱赞助这个项目。虽然他们的决策是相互独立的，但双方的结果都受对方决策的影响。如果只有某个政府决定赞助这项研究，那么另一个国家的人们也可以不花成本而得到科研成果和使用疫苗，可是各个政府的支付只依赖于本国人口所承担的损失和研发费用。

多米尼加和索婆里亚之间的非合作博弈的支付矩阵如表 11—6 所示。每个国家有“研发”和“不研发”这两个策略可以选择，支付是各种策略组合下各国的损失，以十亿美元为单位。直接可以看出这个博弈是个囚徒困境，每个国家都有一个占优策略，即不研发。

表 11—6　　人口相等时的 SANE 疫苗研发博弈的支付结构（十亿美元）

		索婆里亚	
		研发	不研发
多米尼加	研发	−2，−2	−2，0
	不研发	0，−2	−1.6，−1.6

现在假设两国的人口不相等，其中多米尼加有 1.5 亿人口，而索婆里亚只有 0.5 亿人口。因而，如果两个国家的政府都不研发，由于 SANE，多米尼加的损失是 24 亿美元（0.000 5×150 000 000×32 000），而索婆里亚损失只有 8 亿美元（0.000 5×50 000 000×32 000）。调整后的支付矩阵如表 11—7 所示。

表 11—7　　人口不等时的 SANE 疫苗研发博弈的支付结构（十亿美元）

		索婆里亚	
		研发	不研发
多米尼加	研发	−2，−2	−2，0
	不研发	0，−2	−2.4，−0.8

在这个新的博弈中，“不研发”仍是索婆里亚的占优策略，但是此时多米尼加的最优反应却是选择“研发”。是什么改变了多米尼加的选择呢？很明显，在改动后的博弈中，人口的分布是不对称的，多米尼加将要遭受很大一部分疾病导致的损失，以致独自选择研发变得划算了，即使它知道索婆里亚要搭便车分享研发的收益，多米尼加也会如此选择。

表 11—7 所示的研发博弈已经不再是囚徒困境。我们可以看到困境通过规模不对称性解决了，大国选择了领导者的角色，并为整个世界提供了福利。

在国际外交中，常常可见在囚徒困境中出现领导者的局面，领导者的角色通常是落到最大型的或最完备的参与人身上，这种现象叫做“以小欺大”。[7] 例如，在第二次世界大战后的很多年里，美国一直都承担着不成比例的防卫同盟（如 NATO）的开支，即使是在日本和欧洲等贸易伙伴实行贸易保护主义的时候，也保持着相对自由的国际贸易政策。在这些情况下，大型或完备的参与人愿意接受

领导者的角色，是因为它的利益与所有参与人的共同利益紧密地联系在一起；如果大型参与人是整个集团中很大的一个组成部分，这种利益的集中就很明显了，大的参与人在行动中也有望更加合作。

11.5 实验证据

很多人做过囚徒困境的实验[8]，证明博弈中的合作行为可能而且的确会发生，即使在有限次重复博弈中也不例外。很多参与人一开始就选择合作，而且只要对手合作就会继续合作。只是在有限次重复博弈的最后几个回合中，背叛行为才会出现。虽然这样的行为违反了逆向归纳法的逻辑，但只要维持一段期限合理的时间，这样的行为仍是有利可图的，跟一开始就选择背叛比起来，会得到更高的支付。

这样的行为可以从不同角度给予合理化的解释：参与人可能不知道博弈关系何时结束；参与人可能认为合作的声誉会延续到其他博弈中，无论面对的是相同的还是不同的对手；参与人可能认为对方是天真的合作者，并且愿意冒点风险在几回合的博弈中测试这个假设，如果成功，会有很长一段时间的高支付。

在某些实验中，参与人参与多个回合的博弈，每个回合由有限次重复博弈组成，每一回合里的博弈中面对相同的对手，但不同回合的博弈中面对不同的对手，所以有机会在前一回合中学到合作的经验并且在下一回合中采用。这解释了为什么合作在前几回合比后几回合持续得久，也暗示了依据逆推法放弃合作行为是从经验中学来的，因为随着时间流逝，参与人更加了解他们行动的代价。另一个可能是参与人学到了他们要比对手先背叛，所以随着回合的增加首次背叛的时间也更早了。

设想你参与一场囚徒困境博弈，并发现自己采取的是合作模式，而已知的博弈终点正在逼近，那么你将何时选择背叛呢？你不想太早，因为还有很多潜在的利益可图；但也不想太晚，因为你的对手可能早你一步背叛，让你的支付变低。事实上，你何时背叛是不可知的。如果可知的话，你的对手就会在你计划背叛前先下手为强。如果你的背叛策略是未知的，则合作模式的解除必须对于两个参与人来说都包含某些不确定性，例如混合策略。很多薄弱的合作关系，例如罪犯之间、告密者与警察之间，就借助了这样的不确定性。

在许多真实世界的情况以及实验中，都会看到合作关系在重复博弈的最后阶段瓦解，长途自行车赛就是一个例子。比赛的大部分时间都有合作关系，参赛者轮流骑在前面，让后面的人可以顺风地滑流前进。然而，当比赛快到终点的时候，每位参赛者都想着要冲刺。同样，每当大学临近期末的时候，写着“不收支票”的告示牌常常出现在大学城的商店门口。

在两个参与人的囚徒困境中，计算机模拟可以做很简单到很复杂的条件策略实验。最有名的实验是由罗伯特·阿克塞尔罗德（Robert Axelord）在密歇根大学主持的。他邀请大家提供计算机程序，每个程序为重复 200 次的囚徒困境制定

一个策略。有 14 个人参与实验。阿克塞尔罗德举行了一场联盟赛，将这 14 个人分成 7 组，两人一组进行重复 200 次的博弈。比赛的分数都记录下来，然后再重新分组，直到两两都比赛过，最后把各人面对不同对手的分数加起来，看哪个程序最厉害。阿克塞尔罗德惊奇地发现“友好”的程序表现较好，最厉害的前 8 个程序都不是首先背叛的，冠军策略竟然是最简单的策略——以牙还牙策略［由加拿大博弈论学者阿纳托尔（Anatole）提出］。急于背叛的程序提早获得背叛的支付，但之后承受互相背叛的低支付；另一方面，总是友好且合作的程序容易被对手所利用。阿克塞尔罗德这样解释以牙还牙策略成功的原因：即刻原谅、仁慈、反应及时、清楚明白。

按照阿克塞尔罗德的说法，要赢得重复的囚徒困境必须遵守四个简单的原则：“不要嫉妒，不要先背叛，对合作和背叛都进行回报，不要太聪明。”[9]以牙还牙策略体现了这四个理想条件：它不嫉妒；它只从自身的利益出发，不会一直都想着超越对手。另外，以牙还牙策略清楚地记住教训，它不会先背叛，只报复对手前一次的背叛，而且总是以仁慈的手段。最后，以牙还牙策略不会被过分聪明所害，它是简单明确的。事实上，它赢得比赛不是因为帮助参与人获得高支付——这不是赢者通吃的博弈——而是因为拉近了双方的支付；它鼓励合作，同时也避免被利用，这是其他策略所做不到的。

阿克塞尔罗德然后宣布了刚才联盟赛的结果，并举行下一场联盟赛。此时，大家已经有明确的目标，就是设计打败以牙还牙策略的程序。结果胜出的仍是以牙还牙策略！被精心设计用以打败以牙还牙策略的程序不能成功地打败它，并且这些程序互相对抗的结果也不好。阿克塞尔罗德又设计了另一种比赛取代联盟赛。每个程序被复制数份，然后所有程序组成程序库，每种类型的程序以随机方式从程序库里选一个对手跟它对抗。表现较好的程序可以扩充它们在程序库中的比例；而表现差的程序必须减少它们的数量。这是个演化及生物进化的博弈，我们会在第 13 章详加讨论。这里的观点很简单，结果却很吸引人，刚开始狡诈的程序会利用善良的程序，以获得高支付，但随着狡诈程序比例越来越大，每个狡诈程序就越有可能遇到另一个狡诈的程序，然后支付就会降低，数量也会减少，此时以牙还牙策略又开始有好的表现并最终赢得胜利。

然而，以牙还牙策略也有一些不足。最重要的是，它假设在策略执行中不会犯错误。如果参与人本来打算采取合作策略却不小心采取了背叛策略，就可能导致一连串的背叛报复行动，使得两位采取以牙还牙策略的参与人困于坏的结果；此时需要另一个错误来解救他们。当阿克塞尔罗德举行第三场联盟赛，赛中允许若干随机错误时，以牙还牙策略可能被更仁慈的程序打败，这些仁慈程序能容忍偶然的欺骗，看看是错误还是蓄意背叛，当相信这不是错误后才进行报复。[10]

11.6 真实世界的囚徒困境

囚徒困境出现在真实世界的各个不同的领域中，要将所有的情况都加以讨论

是不可能的，但我们借此机会详细讨论四个不同领域的例子。第一个例子考察政府的政策安排；第二个例子讨论劳动仲裁；第三个例子源于生物进化（我们在第13章会进一步讨论）；第四个例子讨论解决价格博弈囚徒困境的途径——“许诺最低价”策略。

11.6.1 政府招商竞争

这个例子源于《时代》（*Time*）杂志中一篇名为《州与州之间的非赢战争》（*A No-Win War Between the States*）的文章。[11]该文章讨论了这样一个事件：每个州都使用财政激励政策来鼓励企业到本州建厂，或使本州的工厂留在本州。这种财政激励主要是采用削减公司税和房产税的形式，提供这种激励的州要付出相当大的代价，但是如果成功的话，收益也会很大。大型制造企业会为本州居民带来增加就业机会的承诺，以及潜在的营业税的增加。然而，提供这种激励的过程却是一个囚徒困境。各州都有提供激励这一占优策略，但是最后均衡的结果对于大家来说都是不利的，反而不提供激励的结果更为有利。

让我们考虑有两个州参加的博弈，我州与他州。两个州目前都有相同数量的厂商，并考虑提供同样的激励措施来吸引另一州厂商来本州设厂。每个州都有两个策略——“提供激励”和“不提供激励”。如果两个州都不提供激励，则两者都不需要付出激励的成本，但不会有新的厂商进入，两个州都保持原有数量的厂商。如果我州提供激励而他州不提供，则我州吸引了厂商进入而他州失去了厂商，所以我州占了他州的便宜（假设我州已经在先前做了充分的调查，激励措施的成本会被新厂商到来带来的收益冲销）；相反，如果他州提供激励而我州不提供，则他州会占我州的便宜。如果双方都提供激励，两个州的厂商会互相流动，但是最后各州厂商的数量会跟没有激励时一样多，此时两个州的结果和原来一样，但付出了高的成本。

给定所有可能的策略和支付，我们构造了一个简单的支付表，如表11—8所示。支付数值代表了结果的等级。当一州提供激励而其对手没有的时候，它得到最高的4，因为该州虽然付出了激励的成本但从新来厂商处获得了利益；而另一方面，没有提供激励的对手得到最低的1，因为该州虽然没有付出激励成本但失去了厂商。在这中间，两个州都没有付出激励成本，双方厂商维持不变，则各自得到3；双方都付出激励成本，双方厂商数量维持不变，则分别得到2。

表11—8　　激励措施提供的囚徒困境的支付表

		他州	
		提供激励	不提供激励
我州	提供激励	2，2	4，1
	不提供激励	1，4	3，3

如表11—8所示的支付表是一个典型的囚徒困境，两个州的占优策略都是提

供激励，纳什均衡使得双方都提供激励。当然，如同所有的囚徒困境一样，两个州如果都不提供激励会有较好的结果。其他政策也有相同的情况。政府在决定是否给国内厂商提供出口补贴的时候也会遇到相同的困境，因为知道其他国家也会同时做出相同的决定。

这样的困境有解吗？从获得激励的厂商的角度来看，这样的均衡结果是好的，但是从政府以及必须承担更重税负的老百姓的角度来说，不提供激励的结果会更好些。由联邦立法来禁止这样的激励措施可以解决这个困境，这属于由第三方惩罚的解决办法那一类。

在欧盟，成员国都通过减税和放松管制来吸引厂商，“向底线冲刺”正在引发类似的担忧，这可能导致更少的政府收入和威胁到各国的社会服务。欧盟需要制定统一的规则来解决这一困境，但至今还没有达成任何协议。

11.6.2 劳动仲裁

第二个例子摘自奥利·阿申弗尔特（Orley Ashenfelter）和戴维·布卢姆（David Bloom）合著的文章《囚徒困境中的魔鬼代言人：律师》（*Lawyers As Agents of the Devil in a Prisoners' Dilemma Game*）[12]，文中作者论证了许多劳方与资方之间的仲裁都具有囚徒困境的结构。对于是否保留仲裁的法律辩论，劳资双方都认为是必要的。在对方不请律师的情况下，自己请律师对仲裁的过程有实质性的帮助。所以，若对手没有找律师的话，你最好去请一个；同时如果对手找了律师的话，你也最好去请一个。因此，聘请律师是占优策略。不幸的是，阿申弗尔特和布卢姆指出都找律师对于仲裁博弈的结果影响很小；也就是说，双方都不请律师和双方都请律师，劳方赢得仲裁的概率都是一样的。当然，请律师要花钱，就像州政府提供的激励措施一样。参与人都想请律师并为此花费，但对结果却没什么帮助。因此，双方都请律师的均衡结果比没请律师的均衡要差，由于双方都要付出相当多的律师费，但这些钱对最后是否赢得仲裁没有影响。

表 11—9 给出了阿申弗尔特和布卢姆的研究结果，表中的数值是在各种不同策略组合下估算出的资方赢得仲裁的概率。劳方聘用律师对付未聘任律师的资方时，比起原先不聘任律师时，赢得仲裁的概率从 56%提高到 77%；反过来，资方也有类似的增加（从 44%到 73%）。但比较双方都聘用律师和双方都不聘用律师时，赢的概率只稍微改变（略升 2%，加上误差）。所以聘用律师的钱对于增加赢得仲裁的概率并没有什么帮助；如果双方都不请律师的话，就可以省下这笔钱。

表 11—9　　在仲裁博弈中估计资方赢的概率

		劳方	
		聘用律师	不聘用律师
资方	聘用律师	46%	73%
	不聘用律师	23%	44%

为找到解决困境的出路，我们再次求助于制定立法惩罚机制的可能——联邦

仲裁法的建立可以排除双方请律师的可能性。在这个博弈中，如果双方保持持续的关系的话，使用条件策略可以延续合作；如果劳资双方一开始就是合作模式，则他们会担心一旦背叛后再也不能回到不请律师时的结果，这足以阻止背叛（请律师）的可能。

11.6.3 生物进化

第三个例子来自生物进化领域，是著名的园丁鸟困境。[13]园丁鸟中的雄鸟在树上建造复杂的鸟巢来吸引雌鸟，而雌鸟对未来伴侣的筑巢很挑剔，由于这个原因，雄鸟通常外出寻找并摧毁其他雄鸟的巢。当它们外出破坏别的鸟巢的时候，也必须承担自己的鸟巢被摧毁的风险。雄鸟间的竞争，也就是掠夺和守卫，这个博弈有着囚徒困境的结构。

鸟类学家构造了两只鸟竞争博弈的支付表。它们有两种策略——“掠夺”或“守卫”，支付表如表 11—10 所示。GG 表示自己和对手都“守卫”时的支付，GM 表示自己“守卫”而对手“掠夺”时的支付。类似地，MM 代表自己“掠夺”而对手也“掠夺”时的支付，MG 代表自己“掠夺”而对手“守卫”时的支付。细致的科学研究发现 MG>GG>MM>GM。换句话说，园丁鸟博弈的支付有着与囚徒困境相同的结构，鸟的占优策略是“掠夺”，但双方“守卫”的结果却比双方“掠夺”的均衡更好。

表 11—10　　　　园丁鸟的困境

		鸟 2	
		掠夺	守卫
鸟 1	掠夺	MM，MM	MG，GM
	守卫	GM，MG	GG，GG

事实上，就鸟而言，任一特定的园丁鸟所运用的策略并不是理性选择过程的结果。在进化博弈中，我们假设这些策略先天就已经被灌注在鸟的身上，而不同的支付代表着具有不同类型策略的鸟繁衍成功的概率。从而此类博弈的均衡就显示了预期可被观察到的鸟群的类型特征。例如，如果“掠夺”是占优策略，则我们看到的鸟都是掠夺者。然而，由于存在囚徒困境，均衡并不是最好的结果。而要解决园丁鸟困境，我们可以诉诸博弈中互动反复进行这一性质。雄鸟在几个繁殖季节面对不同或相同的对手，会让雄鸟依据对手的最后一个行动选择比较灵活的策略。以牙还牙策略等条件策略通常会被进化博弈所采用，以解决困境。我们将在第 13 章中再次涉及进化博弈的思想，并进行更详细的讨论。

11.6.4 许诺最低价

最后，我们回到价格竞争博弈，其中两个商家进行价格竞争，它们使用相同

的价格竞争策略。比如，玩具反斗城和康马特都是全国连锁店，它们经常以广告的形式宣传各自的品牌玩具（和其他的产品）的价格。此外，每家玩具店都有公开的策略：绝对会以市面上看到的最低价格卖出产品（模型和数量必须相同），只要顾客能提供竞争者的印刷广告。[14]

为了讨论的方便，我们假设玩具店对某一玩具只有两种可能的定价（“高价”，“低价”），并且我们假定了各种情况下收益的数值。同时，我们假设玩具反斗城和康马特是某个城市——如蒙大拿州毕林斯市（Billings，Montana）——市场上仅有的两家竞争厂商，以简化分析。

假设两家玩具店博弈的基本结构如表 11—11 所示。如果双方都采取低价，则各得一半的市场并赚得 2 000 美元；如果双方都采取高价，则一样各得一半的市场，但因价格高所以赚得 3 000 美元。最后，如果两家定价不同，则高价者市场收益为 0，而低价者获得 4 000 美元。

表 11—11　　玩具反斗城与康马特的竞价博弈

		康马特	
		低价	高价
玩具反斗城	低价	2 000，2 000	4 000，0
	高价	0，4 000	3 000，3 000

如表 11—11 所示的博弈显然是个囚徒困境，“低价”是双方的占优策略，虽然“高价”能使得它们获利更多。但如同先前所述，玩具店实际采取了第三种价格策略：对它的消费者承诺最低价格。这个策略将如何影响存在于两个玩具店之间的囚徒困境呢？

现在考虑如果允许玩具店选择“高价”、“低价”以及“许诺最低价”，结果将会如何。“许诺最低价”策略让玩具店采取“高价”，但承诺降到跟任何竞争者一样低的价格。如果对手也采取“高价”，则“许诺最低价”会让它从中得到好处；如果对手采取“低价”，它也不会遭受任何损失。我们看到新博弈的支付结构如表 11—12 所示，一方采取“低价”而另一方采取“许诺最低价”策略的组合，跟双方都采取“低价”的组合是等价的；而一方采取“高价”且另一方采取“许诺最低价”策略的组合（或双方都采取“许诺最低价”策略的组合），跟双方都采取“高价”的组合是等价的。

表 11—12　　玩具反斗城与康马特的许诺最低价博弈

		康马特		
		低价	高价	许诺最低价
玩具反斗城	低价	2 000，2 000	4 000，0	2 000，2 000
	高价	0，4 000	3 000，3 000	3 000，3 000
	许诺最低价	2 000，2 000	3 000，3 000	3 000，3 000

使用分析同时行动博弈的标准方法，可知“许诺最低价”弱占优于“高价”，一旦剔除“高价”后，“低价”对“许诺最低价”而言也是弱的劣策略，最后纳什均衡结果是双方都采取“许诺最低价”的策略。均衡时，双方得到 3 000 美元——这个支付与原博弈中双方都采取“高价”时的支付一样。“许诺最低价”策略的出现解决了当玩具店只能选择“高价”和“低价”时的囚徒困境。

这到底是怎么回事呢？“许诺最低价”策略扮演的是惩罚机制的角色。虽然玩具反斗城的“许诺最低价”策略一开始采取“高价”，但它许诺会采用康马特采用的低价，所以大幅降低了康马特采取“低价”的利益。另外，许诺采用康马特采用的低价也会损害玩具反斗城自身的利益，因为它必须接受“低价”带来的低收益。因此，只要有人背叛，“许诺最低价”策略就会惩罚双方。这就像第 11.3 节讨论的帮派老大例子一样；只不过这个惩罚机制——以及它支持的高的均衡价格——能在全国的各个城市的市场上真实见到。

证明这种“许诺最低价”策略的弊端的实验证据是存在的，但是不多。一些研究发现在采取低价销售的市场上也存在这样的策略。[15]然而，这个结果应该让所有消费者提高警惕，虽然商家以竞争的名义采取“许诺最低价”策略，但所有的商家都采取这个策略的最终结果，对于商家而言是更有利的，损害的反而是消费者的利益。

11.7 总　　结

囚徒困境大概是最著名的策略博弈。每个参与人都有占优策略（背叛），但均衡结果对于所有参与人来说都是不利的，反而都采取劣策略（合作）会比较好。最为大家熟知的囚徒困境解决办法是重复博弈。在有限次重复博弈中，未来合作的现值最后会是 0，逆向归纳会得到没有合作行为的均衡；在无限次重复博弈（或不确定何时结束的博弈）中，可以利用适当的条件策略来达成合作，如以牙还牙策略或冷酷策略。无论何种情况，只有当合作的现值超过背叛的现值时，才有合作的可能。更普遍的是，“没有明天”或短期的关系会减少参与人合作的可能。

惩罚机制也可以解决囚徒困境，它是通过改变选择背叛的参与人的支付来实现的。如果有大型或强大的参与人存在，就会有第三种解决办法，因为对于大型的参与人而言，背叛的损失比合作的收益要大。

有实验证据发现，参与人的合作时间往往比理论预期的要持久，这可能是因为参与人对博弈具有不完全知识，或者参与人对合作的价值认识不同。以牙还牙策略被视为是简单、仁慈、反应迅速且善于原谅的策略，平均而言，它在重复博弈中的表现非常好。

囚徒困境在各个领域都有可能发生，几个特别的例子，包括政策安排、劳动仲裁、生物进化以及最低价格竞争，都让我们知道如何用囚徒困境框架来解释实际行为。

关键术语

复利（compound interest）
条件策略（contingent strategy）
贴现因子（discount factor）
有效收益率（effective rate of return）
冷酷策略（grim strategy）
无限期（infinite horizon）
领导（leadership）
惩罚（penalty）
现值（present value，PV）
惩罚（punishment）
重复博弈（repeated play）
以牙还牙策略（Tit for Tat，TFT）
触发策略（trigger strategy）

习　题

1. “如果囚徒困境重复 100 次，则参与人肯定会达到合作的结果。”这句话对还是错？解释并以一个例子说明你的答案。
2. 两个人，贝克和卡特勒，参加一个博弈，她们选择并分配一份奖金。贝克决定奖金大小，她选择 10 美元或 100 美元；卡特勒选择平均分配或是她得到 90%而贝克得到 10%，写下这个博弈的支付表并找出均衡。
   (a) 当她们俩同时决策。
   (b) 当贝克先行。
   (c) 当卡特勒先行。
   (d) 这是不是一个囚徒困境？为什么？
3. 一个公司有两个部门，每个部门都有一个经理。部门经理都按照他们在提高部门生产率上的努力来获取薪酬，而努力的程度是通过与其他经理和其他部门的比较来评价的。如果两个经理都被评为“高努力”，则每人得到 150 000 美元/年；如果两个都被评为“低努力”，则每人只得到 100 000 美元/年。但是如果其中一个经理表现出“高努力”而另一个表现出“低努力”，那么“高努力”的经理得到 150 000 美元的工资外加 50 000 美元的奖金，同时另一个（“低努力”）经理只得到降低了的薪金 80 000 美元（因为在比较评比中较次的工作表现）。经理独立做出努力水平的决策，同时不知道对方的选择。
   (a) 假设付出努力是没有成本的，画出这个博弈的支付表，找出博弈的纳什均衡并且解释这个博弈是否是一个囚徒困境。
   (b) 现在假设付出高努力是需要经理付出成本的（如质量的信号）。特定地，假设“高努力”使得选择它的经理付出等价于 60 000 美元/年的成本，画出这个新博弈的支付表，找出博弈的纳什均衡，解释这个博弈是否是一个囚徒困境并且说明它与（a）中博弈的差别。
   (c) 如果高努力的成本是 80 000 美元/年，那么这与（b）中的博弈又有什么不同？新博弈的均衡是什么？解释这个博弈是否是囚徒困境，并且说明

它与（a）和（b）中博弈的不同。

4. 在一个小镇上，有一群爱吃比萨的居民，但是只有两个比萨店：唐娜的厚薄饼和皮尔斯的比萨，它们对自己的比萨定价。为简化起见，假设它们的策略只有“高价”和“低价”两种。如果定了高价，那么比萨店将得到每块饼 12 美元的利润；如果定了低价，每块饼只能得到 10 美元的利润。每个店都有一群忠实的顾客，无论价格如何，他们每个礼拜能购买 3 000 块饼。还有一个每周 4 000 块饼的流动需求。这些顾客对价格比较敏感，他们只买价格最低的那个店的饼。如果两个店定价一样，那么它们就平分这群顾客的需求。
   (a) 画出这个比萨定价博弈的支付表，支付按每个店的每周利润（千美元）来计量。找出这个博弈的纳什均衡，并且解释为什么这是个囚徒困境。
   (b) 现在假设唐娜的厚薄饼有一个更大的忠实顾客群，他们能许诺每周11 000 块（而不是 3 000 块）的需求，每块饼的利润和流动需求都保持不变。画出这个新博弈的支付表，并且找出纳什均衡。
   (c) 唐娜的厚薄饼的更大的忠实顾客群的存在是如何帮助“解决”比萨店困境的？

5. 考虑第 4 章中的懦夫博弈，并且博弈的支付更一般化（表 4—14 中的 $k=1$）：

		迪恩	
		转向	径直向前
詹姆斯	转向	0，0	−1，k
	径直向前	k，−1	−2，−2

假设博弈在每周六的傍晚重复进行，如果 $k<1$，则两个参与人会持续合作，（转向，转向）使双方受益；如果 $k>1$，则一方“转向”而另一方“径直向前”的合作会使双方受益，他们每周轮流“径直向前”。这两种合作都能持续吗？

6. 一个市议会由三个议员组成，他们每年投票决定是否加薪。加薪需要两张赞成票，每位议员都希望加薪，但却想投反对票，因为这样会让选民有好的印象。每位议员的支付如下：

   加薪通过，自己投反对票：10
   加薪失败，自己投反对票：5
   加薪通过，自己投赞成票：4
   加薪失败，自己投赞成票：0

   投票是同时进行的，画出三维的支付表，证明在纳什均衡中，加薪都是失败的。验证议员间重复的关系如何在下面两个条件下分别确保每年加薪成功：（1）议员每三年一任，每年有一位轮流改选；（2）市民是健忘的，他们只记得今年的投票行为，而以前的就忘记了。

7. 在习题 3 的例子中，两个部门经理关于“高努力”和“低努力”的选择决定了他们的薪金。在该习题的（b）中，假设付出高努力的成本是 60 000 美元/年，我们这里进一步假设习题 3（b）中的博弈会重复许多年。这样的重复允许了一种不常见类型的合作的存在：一个经理选择“高努力”而另一位选择“低努力”，这种合作要求选择“高努力”的经理付出一个单边转移支付给选择“低努力”的经理，从而他们最终的支付是相同的。

(a) 多大的单边转移支付能保证最终两个经理的支付相等？在合作协议下，每个经理每年的支付是多少？

(b) 重复博弈中的合作使得每个经理选择相应的努力水平，同时选择“高努力”的经理付出了指定的单边转移支付，背叛使得他拒绝付出单边转移支付。当收益率满足什么条件时，合作在这个经理人博弈中能持久？

8. 考虑宝宝乐（Child's play）和孩子角（Kid's Korner）间的博弈，它们都是生产并销售木质小孩摇椅的公司。每家公司都可以对一张标准的摇椅定高价或低价。如果它们都采取高价，会分别得到 64 000 美元（每年）的收益；如果一家采取低价而另一家采取高价，采取低价的公司每年获得 72 000 美元，高价公司获得 20 000 美元；如果它们同时都采取低价，则分别得到 57 000 美元的收益。

(a) 证明这个博弈有囚徒困境的结构（比较不同策略组合的支付）。如果博弈只有一个回合，则纳什均衡策略是什么？支付是什么？

(b) 如果两家公司决定参与为期四年的博弈，则每家公司最后的总收益是多少（不计贴现）？解释你是如何得到这个答案的。

(c) 假设两家公司永久参与博弈，让它们都使用冷酷策略，并都采取高价，如果某一方背叛，则剩下的博弈它们都采取低价，背叛那年得到的收益是多少（如果对方采取合作策略）？背叛之后双方在未来每年的损失是多少？如果 $r=0.25(\delta=0.8)$，合作划算吗？找出让两家公司维持合作的 r（或 δ）值的范围。

(d) 假设博弈每年重复进行，如果世界将于 4 年后在无人预见的情况下毁灭，则博弈最后每家公司的总收益是多少（不计贴现）？把答案和（b）做比较，这两个答案一致么？解释为什么。

(e) 假设两家公司知道每年有 10％的概率它们中的某一家会破产。如果破产，则关系就结束。如果 $r=0.25$，这个假设会不会改变公司的行动？如果破产概率上升到 35％的话，结果又如何呢？

9. 你要决定是否将 100 美元投到你的一个朋友的企业，一年的时间这笔投资可以升值为 130 美元。你同意你的朋友还给你 120 美元，同时保留 10 美元给他自己。然而，也有可能你的朋友将全部的 130 美元卷款逃跑。任何你不投资到你朋友企业的资金都可以进行安全的投资得到一般的利率 r，来年可得到 $100(1+r)$ 美元。

(a) 画出这种情况下的博弈树，用逆向归纳法求解均衡。

下面设想这个博弈进行无穷多次，每年你都可以对你朋友的企业投资 100 美元，分配投资收益的 130 美元的协议同上不变。从第二年开始，你可以依据前些年份你朋友履行协议的情况来决定是否继续投资。两个相邻年份间的利率是 r，与在外面投资的收益率相同，并且对你和你朋友来说都是一致的。

(b) 在 r 取什么值的时候，重复博弈的均衡结果是你每年都投资你朋友的企业而他都履行协议付款？

(c) 如果利率是每年 10％，有没有其他的什么投资收益分配方案成为无限次重复博弈的均衡结果，使得每期你都投资你朋友的企业而他都按协议付款？

10. 考虑习题 4 中介绍的比萨店：唐娜的厚薄饼和皮尔斯的比萨，设想它们不是只选择两个可能的价格，而是可以选择一个特定价格来使得利润最大化。假设生产一块比萨的成本是 3 美元（对于每家店来说），经验和市场调查显示每家店的销量（Q）和价格（P）之间的关系是

$$Q_{唐娜}=12-P_{唐娜}+0.5P_{皮尔斯}$$
$$Q_{皮尔斯}=12-P_{皮尔斯}+0.5P_{唐娜}$$

则每家店每个礼拜的利润（Y，千美元）分别是

$$\begin{aligned}Y_{皮尔斯}&=(P_{皮尔斯}-3)Q_{皮尔斯}\\&=(P_{皮尔斯}-3)(12-P_{皮尔斯}+0.5P_{唐娜})\end{aligned}$$
$$\begin{aligned}Y_{唐娜}&=(P_{唐娜}-3)Q_{唐娜}\\&=(P_{唐娜}-3)(12-P_{唐娜}+0.5P_{皮尔斯})\end{aligned}$$

(a) 像在第 5 章一样，用这些利润函数求出每家店的最优反应规则，并且用这些最优反应规则找出这个定价博弈的纳什均衡。每家店在均衡时都定什么价格？每家店每周的利润是多少？

(b) 如果两家店一起经营，选择共同最优的价格 P，那么每家店的利润是

$$\begin{aligned}Y_{唐娜}=Y_{皮尔斯}&=(P-3)(12-P+0.5P)\\&=(P-3)(12-0.5P)\end{aligned}$$

定什么价格可使得共同的利润最大？

(c) 设想这两家店之间的博弈重复进行，力图保持（b）中计算出来的使共同利润最大化的价格。它们每个月都印出新菜单并在整个月中都遵守这个价格；任何一个月，它们都可以背叛合作协议。如果它们中的一家遵守协议，那么另一家店选择背叛的最优价格是多少？它的背叛利润是多少？通过采用冷酷策略，贴现因子是多少的时候共谋可以保持？

11. 考虑以下由密歇根大学的詹姆斯 · 安德烈奥尼（James Andreoni）和哈尔 · 瓦里安（Hal Varian）提出的博弈。[16] 博弈中有个中立的裁判，有两个参与人——行和列。裁判给每个人两张卡片：2 和 7 给行，4 和 8 给列。这是两个人的共同知识。然后博弈同时且独立进行，每个参与人要给裁判他的某张卡片，裁判根据收集到的卡片上的数字给予支付——该支付来源于公共基金而不是参与人的口袋——并以美元计算。如果行选择是数字小的卡片 2，则他得到 2 美元；如果行选择的是数字大的卡片 7，则列得到 7 美元。如果列选择的是数字小的卡片 4，则他得到 4 美元；如果列选择的是数字大的卡片 8，则行得到 8 美元。

(a) 说明完整支付表如下表所示。

		列	
		低	高
行	低	2，4	10，0
	高	0，11	8，7

(b) 纳什均衡是什么？证明这个博弈是个囚徒困境。

现在假设这个博弈有两个阶段，裁判跟以往一样发卡片，每个人得到的全部卡片的数值是共同知识。在第一阶段，每个参与人从他口袋中拿出一笔钱给裁判，裁判把这些钱存入一个托管账户，这笔钱可以为零但不能为负。这些钱的处理规则如下，并且是共同知识：如果列选择数字大的卡片，则裁判把行存在托管账户中的钱转给列；如果列选择数字小的卡片，则裁判把行在该账户中的钱还给行。行的卡片选择也是同样的道理。在第二阶段，每个参与人把他的某张卡片交给裁判，然后裁判依据支付表从公共基金中拿钱给参与人，并根据上述规则处理他们在托管账户中的钱。

(c) 找出此两阶段博弈的逆向归纳均衡（子博弈完美均衡），这能解决囚徒困境吗？托管账户起了什么作用？

【注释】

[1] 触发策略要求以背叛作为报复手段，通常被称为惩罚，以与原有的偏离合作的决定相区别。

[2] 本章附录会详细讨论无穷级数的求和问题。

[3] R 也可以用 r 和 p 来表示，即 $R=(1+r)/p-1$。

[4] 如果信息传导不完全，背叛相关的成本可能就比较小，正如当参与人很多时，识别背叛者和一起协作进行惩罚就变得比较困难。同理，如果对手不能迅速地识别背叛，背叛的收益可能会比较大。

[5] 注意，我们曾在 11.2 节的重复博弈例子中得到了相同的结果。

[6] 有点像在电影《沉睡者》(*Sleeper*) 中的瑞普·凡·温克 (Rip Van Winkle) 或者伍迪·艾伦 (Woody Allen)，但时间要短很多。

[7] Mancur Olson, *The Logic of Collective Action*, Cambridge: Harvard University Press, 1965, p. 29.

[8] 关于囚徒困境的实验文献相当多，想了解概要介绍可以参考 Alvin Roth, *The Handbook of Experimental Economics*, Princeton: Princeton University Press, 1995, pp. 26-28。心理学和经济学期刊也可以提供更多的参考。我们描述的一些例子可以参考 Kenneth Terhune, "Motives, Situation and Interpersonal Conflict within Prisoner's dilemmas," *Journal of Personality and Social Psychology Monograph Supplement*, vol. 8, no, 3, 1968, pp. 1-24; R. Selten and R. Stoecker, "End Behavior in Sequence of Finite Prisoner's Dilemma Supergames," *Journal of Economic Behavior and Organization*, *vol.* 7, 1986, pp. 47-70。罗伯特·阿克塞尔罗德 (Robert Axelrod) 的著作《合作的进化》(*Evolution of Cooperation*, New York: Basic Books, 1984) 则利用电脑模拟，提供了无限次重复博弈的最佳策略。

[9] Axelrod, *Evolution of Cooperation*, p. 110.

[10] 如果要从生物学观点来描述并分析阿克塞尔罗德的计算机模拟，请参考 Matt Ridley, *The Origins of Virtue*, New York: Penguin Books, 1997, pp. 61, 75; 如果要讨论计算机模拟与人类实验的不同，请参考 John K. Kagel and Alvin E. Roth, *Handbook of Experimental Economics*, Princeton, N. J.: Princeton University Press, 1995, p. 29。

[11] John Greenwald, "A No-Win War Between the States," *Time*, April 8, 1995。关于这个话题在国家和地方公共财政领域有很多继续的讨论，如 Stephen Ellis and Cynthi-

a Rogers, "Local Economic Development As a Prisoners' Dilemma: The Role of Business Climate," *Review of Regional Studies*, vol. 30, no. 3, Winter 2000, pp. 315 - 330。

[12] Orley Ashenfelter and David Bloom, "Lawyers As Agents of the Devil in a Prisoners' Dilemma Game," National Bureau of Economic Research Working Paper, No. 4447, September 1993.

[13] Larry Conik, "Science Classic: The Bowerbird's Dilemma," *Discover*, October, 1994.

[14] 玩具反斗城（Toys "R" Us）的价格竞争策略在每间店都看得到告示，而康马特（Kmart）也经过电话证实有相同的策略。类似的策略会出现在许多行业中，包括如在信用卡行业看到的"利率竞争"，请见 Aaron S. Edlin, "Do Guaranteed-Low-Price Policies Guarantee High Prices, and Can Antitrust Rise to the Challenge?" *Harvard Law Review*, *vol.* 111, no. 2, December 1997, pp. 529 - 575。

[15] J. D. Hess 和 Eitan Gerstner 给出了许诺最低价策略导致价格上升的证据，参见 "Price-Matching Policies: An Empirical Case," *Managerial and Decision Economics*, vol. 12, 1991, pp. 305 - 315。Arbatskaya, Hviid 和 Shaffer 给出了相反的证据，他们发现许诺最低价策略导致价格更低，参见 Maria Arbatskaya, Morten Hviid and Greg Shaffer, "Promises to Match or Beat the Competition: Evidence from Retail Tire Prices," *Advances in Applied Microeconomics*, vol. 8: *Oligopoly*, New York: JAI Press, 1999, pp. 123 -138。

[16] James Andreoni and Hal Varian, "Preplay Contacting in the Prisoner's Dilemma," *Proceedings of the National Academy of Sciences*, vol. 96, no. 19, September 14, 1999, pp. 10933 - 10938.

附录　无穷级数求和

现值的计算涉及未来支付的金钱的现在价值的决定，如 11.2 节所述，n 个月后付给我们的金钱 x 的现值只有 $x/(1+r)^n$，其中 r 是每月收益率。但是如果未来每月付给一笔钱 x，则这些钱的现值计算就比较麻烦了。如果这样的支付一直持续到永远，则我们要计算的现值就没有一个确定的尽头，那么这样的计算就需要用无穷级数的思想。

考虑一个囚徒困境中的参与人，他在选择背叛当月获得 36 美元，但未来每个月都失去 36 美元，因为他选择继续背叛而他的对手（采用以牙还牙策略）惩罚他。在惩罚的第一个月——发生损失的第一个月或要进行折现的第一个月——他损失的现值是 $36/(1+r)$；在惩罚的第二个月，他损失的现值是 $36/(1+r)^2$；在惩罚的第三个月，他损失的现值是 $36/(1+r)^3$；在惩罚的第 n 个月，他因背叛而损失的现值是 $36/(1+r)^n$。

我们可以把未来损失的所有现值写出来，会有无限多项：

$$PV=\frac{36}{1+r}+\frac{36}{(1+r)^2}+\frac{36}{(1+r)^3}+\frac{36}{(1+r)^4}+\frac{36}{(1+r)^5}+\frac{36}{(1+r)^6}+\cdots$$

或者我们可以用求和的符号来简化表示上式：

$$PV=\sum_{n=1}^{\infty}\frac{36}{(1+r)^n}$$

这个表达式跟前一个等价，读做“从 n 等于 1 到 n 等于无穷大，36 除以（$1+r$）的 n 次方的总和”，因为 36 是共同因子——它出现在每一项中，可以从式子中提到前面，所以原式变成

$$PV = 36 \cdot \sum_{n=1}^{\infty} \frac{1}{(1+r)^n}$$

我们现在要求出这个总和，以计算实际现值。首先把符号化简，以贴现因子 δ 代替 $1/(1+r)$，然后我们要求的和变为

$$\sum_{n=1}^{\infty} \delta^n$$

这里注意 $\delta=1/(1+r)<1$，因为 r 是正数。熟悉无穷级数求和的人，一看这个式子就知道总和会收敛到一个有限值 $\delta/(1-\delta)$。[1] 因为是对小于 1 的数求次方，所以一定会收敛，这里 δ 的 n 次方越来越小，当 n 接近无穷大的时候，它也接近于 0。所以在我们计算的现值中，越后面的项会越来越小，小到总和接近于某个特定值（技术上总是不会达到这个值）。解出总和收敛于 $\delta/(1-\delta)$ 需要复杂的数学，但验证这是正确的答案却是很直接的。

我们用简单的技巧来验证这个结论，考虑前 m 项的和 S_m：

$$S_m = \sum_{n=1}^{m} \delta^n = \delta + \delta^2 + \delta^3 + \cdots + \delta^{m-1} + \delta^m$$

上式两边同乘以（$1-\delta$）得

$$\begin{aligned}(1-\delta)S_m &= \delta + \delta^2 + \delta^3 + \cdots + \delta^{m-1} + \delta^m \\ &\quad - \delta^2 - \delta^3 - \delta^4 \cdots - \delta^m - \delta^{m+1} \\ &= \delta - \delta^{m+1}\end{aligned}$$

两边同除（$1-\delta$）得到

$$S_m = \frac{\delta - \delta^{m+1}}{1-\delta}$$

最后我们令 m 趋于无穷大看上面这个和的极限，以计算原来的无穷级数的和。当 m 趋于无穷大时，δ^{m+1} 的值就趋于 0，因为对一个小于 1 的数求非常大且增加的次方，结果是很小的非负数。因此，随着 m 接近无穷大，前面方程右边化为 $\delta/(1-\delta)$，也就是 m 趋于无穷大时 S_m 的极限。这样我们就证明了无穷级数的和。

我们现在只要代入 r 就可以得到囚徒困境中计算现值的答案，因为 $\delta=1/(1+r)$，所以：

$$\frac{\delta}{1-\delta} = \frac{1/(1+r)}{r/(1+r)} = \frac{1}{r}$$

所以，从下个月开始，未来每个月惩罚的 36 美元的无限个现值总和是

$$36 \cdot \sum_{n=1}^{\infty} \frac{1}{(1+r)^n} = \frac{36}{r}$$

在 11.2 节中，我们可以用这个值决定参与人是否应该永久背叛。注意，如果我们把未来持续的概率 $p \leqslant 1$ 考虑进来，不会改变无穷级数的运算过程，我们仅需将 r 替换为 R，将贴现因子 δ 替换为 $p\delta$ 即可。

记住你只需要计算未来损失的现值，今天损失的 36 美元的现值还是 36 美元。所以如果你要找出从今天开始每月损失 36 美元的现值，你只需要把本月损失 36 美元加上下一个月开始损失的现值，这个现值就是我们刚才计算过的 $36/r$。因此，加上 36 美元后是 $36+36/r$，或 $36[(r+1)/r]$，等于 $36/(1-\delta)$。同理，如果你要计算参与人在囚徒困境中某条件策略下的一系列收益的现值，你不用把首期的收益贴现，你只要把未来的收益贴现即可。

【注释】

[1] 如果无穷级数的和接近某个特定值，我们称为收敛，随着项数增加而更接近这个值。如果无穷级数的和越加越大，我们称为发散。无穷级数收敛要求后面的各项越来越小。

第 12 章　集体行动博弈

前面各章节谈到的博弈和策略局势通常只包括两个或三个参与人的互动。此类博弈在学术、商业、政治及日常生活中屡见不鲜，因而有研究了解的必要。但是，许多经济、社会和政治互动会同时包含大量的参与人，例如职业规划、投资计划、高峰期的行车路线，甚至做研究，都有可能因为别人的行动而使自己获益或付出代价。如果你曾身处这种情形，你可能会觉得某些事不对劲——太多的学生、投资者和通勤者拥挤在你想去的地方。如果你试图组织同学或朋友去做某件事，你可能会感到沮丧——自愿者太少了。换句话说，社会中的多人博弈的结果通常不能令社会全部成员甚至大多数成员感到满意。在本章，我们将以我们已经在前面章节建立起的博弈论观点来研究这类博弈，我们会让你明白出了什么问题，又如何去处理。

一般形式下，这种多人博弈需要考虑**集体行动**（collective action）问题。社会和集体的最优目标可以通过社会成员采取某些特定的行动来达到，但是这些行动对于个体成员来说却不符合其最优利益。换言之，社会最优结果并不会像博弈的纳什均衡那样自动达成。因此我们需要考虑如何修正博弈以达到最优结果，或至少可以改善现有的不令人满意的纳什均衡。要这么做首先得了解此类博弈的性质。我们发现它们有三种形式，这三种形式大家都已熟悉：囚徒困境、懦夫博弈和安全博弈。虽然本章的重点是同时有多个参与人的博弈，但仍不妨从有两个参与人的博弈开始。

12.1 双人集体行动博弈

假如你和你的邻居都是农夫，建设一项灌溉工程对你们都有利。你们可以一同参与这项工程，也可独自建设。不过，一方将工程建设完成后，则另一方也可自动获益。因此每个人都想把建设工作留给对方。这是策略互动的本质，也是集体行动的障碍。

在第 4 章，我们曾遇到这样的博弈：三位邻居决定是否建造花园，而花园可以共享，结果三人都不想建，这是囚徒困境式的问题。这里，我们将用更具一般性的支付结构来分析。同时，在花园博弈中我们把结果分成六个等级；当我们描述更一般的博弈时，我们必须考虑每个参与人成本和收益的一般形式。

灌溉工程有两个重要特征：其一，其好处是**非排他性的**（nonexcludable），即没有贡献的人也会受益；其二，其好处是**非竞争性的**（nonrival），即你的利益不会因别人也获益而减少。经济学家把这样的项目称为**纯公共物品**（pure public good），国防是经常被提到的例子。相反，**纯私人物品**（pure private good）是绝对排他性的、竞争性的——没有贡献就得不到利益，一旦有人得到利益则其他人就得不到，面包是纯私人物品的好例子。大部分物品处于这两种极端物品的中间，具有某种程度的排他性及竞争性。我们无须对此深入探究，但稍微提一下对你学习其他课程和教材会有所帮助。[1]

灌溉工程的成本和收益取决于哪些参与人参与。若你们两人各自单独建造则每人需要花 7 周，若两人一起建造则每人只需要花 4 周。两人共同建造的灌溉工程有较好的质量。单独建造每人得到价值 6 周的收益，一起建造则每人得到价值 8 周的收益。

这个博弈中，参与人必须决定是参与建造工程还是不参与，即偷懒。（可能有一份临时的工作必须要做，但是在快要结束时假装家里有急事提前溜掉。当然你的邻居也可以这样做。）表 12—1 给出了博弈的支付表，数字表示以星期数来衡量的价值。给定表中的支付结构，你邻居不建造时你的最优反应也是不建造：你独自完成工程的收益（6）比成本（7）要少，不建造会得到 0。同理，如果你的邻居选择建造，你不花任何代价就可获益（6），比你一起建造花成本（4）而获益（8）要好。在这种情况下，你就成了邻居努力建造下的**搭便车者**（free rider）。这个博弈的特性是无论你邻居怎么做，你最好都不要参与；对于你的邻居，道理也是一样的。因而，不建造是每个人的占优策略。但若两人一起建造各获得

表 12—1　　囚徒困境式集体行动博弈：版本 1

		邻居	
		建造	不建造
你	建造	4，4	−1，6
	不建造	6，−1	0，0

支付 4 比两人都不建造各获得支付 0 显然更好。所以，这个博弈是一个囚徒困境。

在囚徒困境中，我们可发现集体行动的主要障碍：个人的最优选择（本例中无论另一农夫怎么做，我就是不建造）从社会整体角度来看可能并非最优的，即使这个社会只由两个农夫组成。集体行动博弈中，“社会”的最优结果是指参与人的总支付最大化。在此囚徒困境中，社会最优结果是（建造，建造）。然而，参与人纳什均衡行为并不总是导致社会最优产出，所以集体行动博弈的研究焦点在于改善已知的纳什均衡行为，使结果向社会最优结果移动。我们将会发现，纳什均衡与社会最优结果的背离会出现在各种版本的集体行动博弈中。

现在考虑将支付数字稍加改变，博弈会变成怎样？假设两人都建造的支付不比单人建造的收益好多少：每个农夫获得价值 6.3 周的利益。所以双人建造时每人利润为 6.3－4＝2.3，支付表如表 12—2 所示。博弈仍然是囚徒困境式的，均衡是（不建造，不建造）。然而，当双方一起建造，总支付只有 4.6；最大支付是一方建造而另一方不建造时，此时的总支付为 6＋(－1)＝5。有两种可能的方式达到这一结果。为达到社会最优结果就产生了新问题：谁应当承担－1 的代价来建造，而让另一方搭便车得到支付 6？

表 12—2　　囚徒困境式集体行动博弈：版本 2

		邻居	
		建造	不建造
你	建造	2.3，2.3	－1，6
	不建造	6，－1	0，0

我们还可以变动一下表 12—1 中的囚徒困境式博弈中的数字，从而改变博弈的性质。假设建造的成本下降了，使得给定邻居不建造时则自己选择单独建造是更好的。具体说来，假设单人建造需要 4 周而双人建造则各自只需要 3 周。收益则跟以前一样。表 12—3 表示了变化后的支付矩阵。现在你的最优反应是，若邻居建造则自己就偷懒，如果邻居偷懒则自己就建造。正式地说，这是懦夫博弈，不建是“径直向前”策略（勇敢或不合作），而建造是“转向”策略（妥协或合作）。

表 12—3　　懦夫博弈式集体行动博弈：版本 1

		邻居	
		建造	不建造
你	建造	5，5	2，6
	不建造	6，2	0，0

如果博弈的结果是某一个纯策略均衡，则两者支付总和是 8，这比双方都建造的总支付要小。也就是纳什均衡策略向社会提供的收益没有比双方一同建造的结果来得好。如果懦夫博弈的结果是混合策略均衡，则结果更糟糕，两者的支付相加小于 8。

懦夫博弈式集体行动还有另外一种可能结构，只要对工程的收益做些改变即可。如同囚徒困境第 2 个版本，假设两人共同建造的工程并不比单人建造的好多少。于是每个农夫从共同建造工程中获益仅为 6.3，仍然从单独建造工程中获益 6，请大家自己练习建立这个博弈的支付表。你会发现这仍是懦夫博弈——称懦夫博弈版本 2。两个纯策略纳什均衡还是一样，只有一个农夫去建造，但双方同时建造的总支付只有 6.6，而单独建造的总支付是 8。社会最优结果是单人建造，但每个农夫都偏好他人建造的均衡，所以双方都在等待别人动手建造。或者农夫可以使用混合策略，但其期望支付会更低。

我们之前假设该工程对你和邻居双方都同等有利。但如果一方的建造会伤害另一方又会怎样呢？例如避免自己农田遭受水淹的方法就是将洪水引到他人的农田，此时，若一方建造则另一方的支付将为负数。这是懦夫博弈的第三个版本：别人不建造则我建造，但是都不建造的结果对于集体来说是更好的。

最后，我们还要改变一次囚徒困境的支付，保持双人建造的收益和成本不变，单独建造的收益减少为 3，这会减少搭便车者的收益。此时，如果邻居选择建造，则你的最优反应也是建造。表 12—4 展示了该博弈的支付表。博弈就变成了有两个纯策略均衡的安全博弈：要么都参与，要么都退出。

表 12—4　　安全博弈式集体行动博弈

		邻居	
		建造	不建造
你	建造	4，4	−4，3
	不建造	3，−4	0，0

如同懦夫博弈版本 2，这里的社会最优结果是某一个纳什均衡。但也有一个不同之处，在懦夫博弈版本 2 中参与人对两个均衡的偏好不同，但在安全博弈中他们有相同的偏好，所以安全博弈比懦夫博弈更容易达成社会最优结果。

以上这些例子指出的问题人们早已熟悉，处理这些问题的不同方法所遵循的一般原理也在早先的章节做过讨论。在转向解决方案之前，让我们先看看现实中多人同时互动博弈中上述问题如何显现出来的。

12.2　大群体中的集体行动问题

本节我们扩展一下灌溉工程的例子。考虑有 N 个农民必须各自决定是否参与工程建造。如果有 n 人参与工程建造，每人付出的代价 c 依 n 而定，记为函数 $c(n)$。而且无论是否参与建造，每个参与人都会从工程中获益，该获益是 n 的函数，记为 $b(n)$。因此每位建造者的支付为 $p(n)=b(n)-c(n)$，不参与建造者的支付 $s(n)=b(n)$。

假设你正在考虑是否加入，你的决定将取决于其他（$N-1$）个农夫的做法。

通常，在 n 个人参与而（$N-1-n$）个人退出的情况下，如果你决定退出，参与者的数量仍为 n，而你得到支付为 $s(n)$；如果你决定加入，则参与者数量变为 $n+1$，而你将得到 $p(n+1)$。因此你的最终决策取决于两种情况的支付比较，若 $p(n+1)>s(n)$ 你就参与，若 $p(n+1)<s(n)$ 你就退出。这个规则对于前一节所有版本的集体行动博弈都成立；不同的博弈有不同的行为是因为支付结构有所改变，影响了 $p(n+1)$ 和 $s(n)$。

我们可将 12.1 节中双人集体行动博弈的例子纳入到现在这个更为一般的框架中。若仅有两人，$p(2)$ 就是对方建造自己也建造时自己获得的支付，$s(1)$ 就是对方建造而自己不建造时自己获得的支付，依此类推。从而我们可以将表 12—1 到表 12—4 的支付表写成一般化的代数形式。这种一般化的支付结构见表 12—5。

表 12—5　双人集体行动博弈的一般形式

		邻居	
		建造	不建造
你	建造	$p(2)$，$p(2)$	$p(1)$，$s(1)$
	不建造	$s(1)$，$p(1)$	$s(0)$，$s(0)$

若如下所有不等式成立，则表 12—5 表示的博弈将是一个囚徒困境问题：

$$p(2)<s(1),\quad p(1)<s(0),\quad p(2)>s(0)$$

第一个不等式是说，如果对手选择建造策略，则己方的最优反应策略是不建造；第二个不等式是说，如果对手选择不建造策略，则己方的最优反应策略也是不建造；第三个不等式是说，（建造，建造）的联合支付比（不建造，不建造）高。如果 $2p(2)>p(1)+s(1)$，这就是第一类困境，都参与建造的总支付比只有一人建造的总支付要大。大家可以建立起类似的不等式组来考察 12.1 节中其他类型博弈的支付。

现在我们转向有 n 个参与人的多人集体行动博弈。我们可用支付函数 $p(n)$ 和 $s(n)$ 来构造第三个函数，用以表示社会的总支付，记为 $T(n)$，也是 n 的函数。社会总支付由 n 个支付为 $p(n)$ 的参与者和（$N-n$）个支付为 $s(n)$ 的退出者的支付总和组成：

$$T(n)=np(n)+(N-n)s(n)$$

现在我们要问：参与和退出的人数如何配置可使得总支付 $T(n)$ 最大？让我们改写 $T(n)$ 以便更好地理解这个问题：

$$T(n)=Ns(n)-n[s(n)-p(n)]$$

上述总支付的推导表明，我们可这样计算总支付，即把退出者的支付分配给 N 个人中的每个人，然后从 n 个参与者中扣除退出者的额外的利益 $[s(n)-p(n)]$。通常 n 增加 $s(n)$ 就会增加，所以第一项 $Ns(n)$ 也会随着 n 增加而增加。若第二项随 n 增加的速度很慢——也就是 $[s(n)-p(n)]$ 很小且不增加，则 $T(n)$ 主要受

第一项影响。$T(n)$ 随 n 的增加而稳定增加，在 $n=N$ 时为最大。这样的情况下将没有退出者。但是，更常见的是 $T(n)$ 会在某个小于 N 的 n 值处取得最大值，即社会最大的总支付可以允许某些人退出建造。在双人博弈中我们已看到这样的例子，下面在多人博弈中我们会见到更多的例子。

现在我们用 $p(n+1)$ 和 $s(n)$ 的函数图形来帮助确定博弈的类别、纳什均衡及社会最优结果。图上的横轴是 n 从 0 到 $(N-1)$，纵轴是支付。两条分开的曲线分别表示 $p(n+1)$ 和 $s(n)$。实际上 n 只取正整数，所以函数 $p(n+1)$ 和 $s(n)$ 应该是离散的点的集合而不是连续的平滑线。但是当 N 很大时，这些离散的点靠得相当近，我们可以把它视为连续的线。为简化起见，我们把 $p(n+1)$ 和 $s(n)$ 视为直线，以后再讨论更复杂的情况。

我们的第一张图（图 12—1）表示 $s(n)$ 完全在 $p(n+1)$ 之上的情况。这是囚徒困境，因为无论多少人参加（也就是不论 n 有多大），退出比参与所得的支付更大；退出是你的占优策略，所以博弈均衡是每个人都退出。但是，两条线都随 n 的增加而上升，无论采取什么行动，参与者越多则你得到的支付就越高。$s(n)$曲线的左截距比 $p(n+1)$ 的右截距要低，也就是 $s(0)<p(N)$，这表示每个人都参与建造时你得到的支付超过每个人都退出建造时你得到的支付。所以大家都参与，每个人的支付会比纳什均衡的支付高。这使得博弈成为囚徒困境。

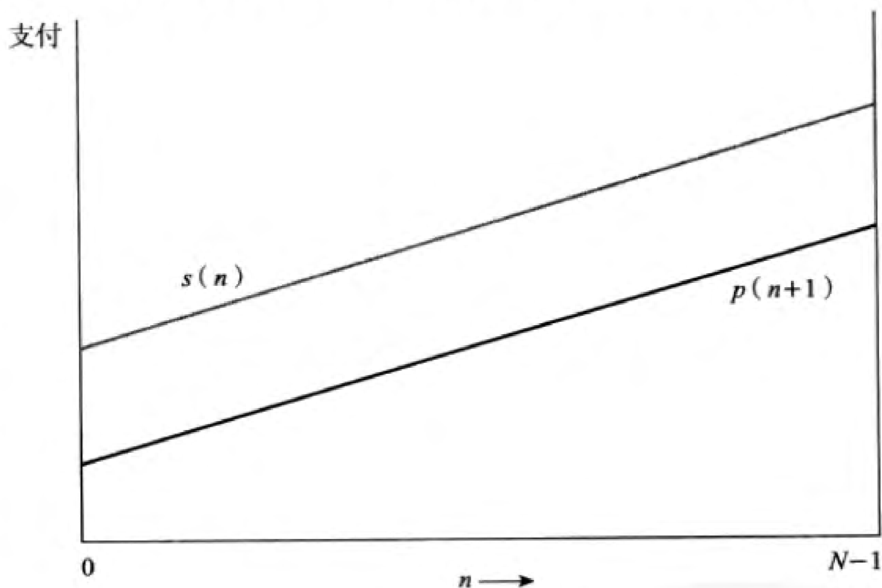


图 12—1　多人囚徒困境

然而，正如我们先前所见，大家都参加建造可使每个人处境变好，但这一事实并不意味着全部人都参与对社会也是最有利的。若囚徒困境中 n 足够大，总支付最大是不可能自行达成的。事实上，如果 $s(n)$ 和 $p(n)$ 之间的缺口随 n 的增加而加速扩大，则 n 向 N 逼近时，$T(n)$ 第二项的负效应会超过第一项的正效应，这时候有必要让一些人退出——也就是，实现社会最优结果时的 n 比 N 要小。当然，这样的结果造成了支付的不公平——未出力者比参与者获益更多，这就增加了克服困境的难度。

图 12—2 是懦夫博弈的例子。当 n 很小时，$p(n+1)>s(n)$，所以如果很少

人参与则你应当参与；若 n 很大，$p(n+1)<s(n)$，所以如果很多人参与则你应当退出。这样的陈述等同于双人博弈中的“你的邻居建造你就偷懒，你的邻居偷懒你就建造”。若两条线相交于某一点，则其对应的 n 的整数值就是纳什均衡的参与者人数。若不如此，则严格讲来该博弈就没有纳什均衡。不过在现实中，如果目前的 n 值是交点左边靠近交点的整数，则会有一个人想参与；若目前的 n 是交点右边最靠近交点的整数，则会有一人想退出。因而参与者的数量会在这一交点附近的小范围内，而我们可大致说这一点是均衡点。

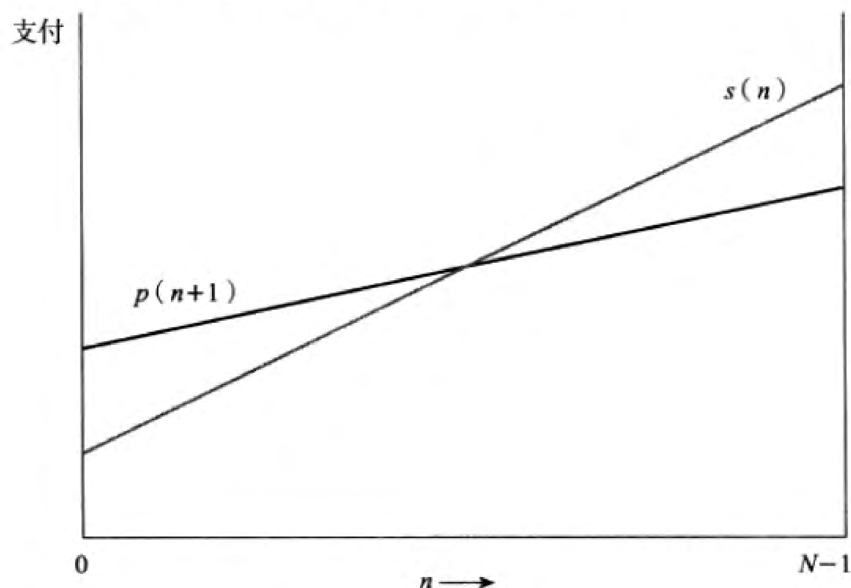


图 12—2　多人懦夫博弈

这里，图 12—2 表明两条线均有正斜率，尽管它们并不需要必定如此；因为随着参与者越来越多，有可能每个人得到的利益会越来越少。这个博弈有一个重要特性，就是当少数人采取行动则任何一个人最好是采取该行动；当多数人采取该行动时，则任何一个人最好是采取另外的行动。

懦夫博弈式的集体行动博弈中，何为社会最优结果？若每位参与者的支付 $p(n)$ 随参与者数量增加而增加，且每一位退出者的支付 $s(n)$ 不会变得显著比参与者的支付 $p(n)$ 大，则所有人参与将会使总支付最大；但一般来说最好让某些人退出。这正是 12.1 节数字例子中，懦夫博弈版本 1 与版本 2 的差异。正如我们在此前指出的，囚徒困境中社会最优时既有参与者也有不参与者，支付在两者之间的分配是不公平的，这使得最优结果难以达到。此类博弈的更一般的例子中，最优的参与者人数有可能比纳什均衡时还要少。我们将在 12.5 节更详细地讨论所有版本博弈的社会最优问题。

图 12—3 表示的是安全博弈。n 很小时，$s(n)>p(n+1)$，所以若很少有人参与，则你应该退出；而 n 很大时，$p(n+1)>s(n)$，所以若很多人参与则你应当参与。这个博弈有两个纳什均衡，大家都退出或大家都参与。如果曲线是上升的——越多人参与则大家的支付就越高，那么很明显右边极端均衡是社会最优结果。问题在于如何达到这一结果。

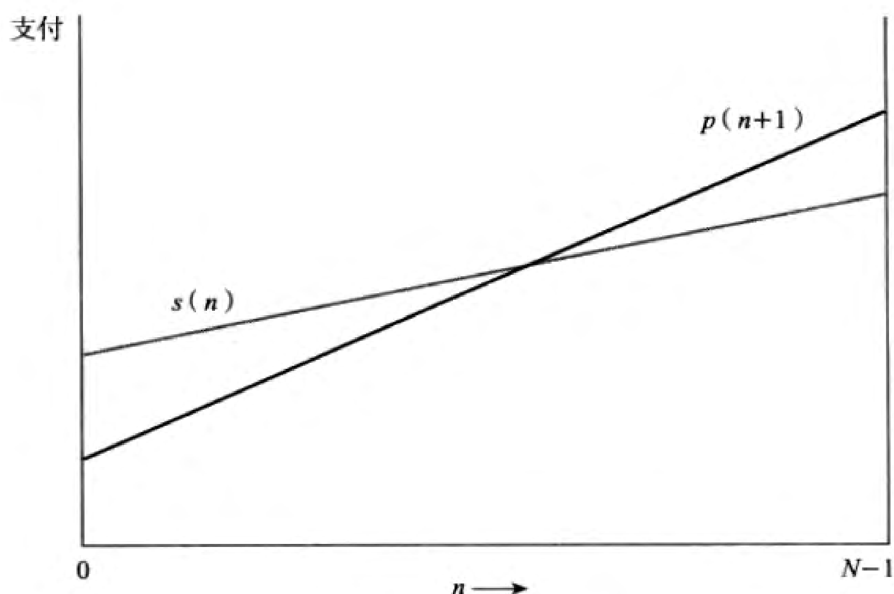


图 12—3　多人安全博弈

当群体中人数 N 非常大时，任何人对博弈都不会有太大的影响，从而 $p(n+1)$几乎跟 $p(n)$ 一样大。这样一来只要 $p(n)<s(n)$则任何人都会退出，以收益跟成本函数表示，亦即 $p(n)=b(n)-c(n)$且 $s(n)=b(n)$。我们发现 $p(n)$ 会绝对小于 $s(n)$，所以只要 N 很大，每个人都会想退出。这就是为什么大群体中公共工程的集体供给问题总是被视为囚徒困境。但如同 12.1 节讨论的双人博弈，这种结果对于小团体不一定适用。而且，集体行动问题涵盖范围很广，而不仅仅局限于是否参与某项未参与者可同样受益的公共工程的问题。12.5 节中我们会见到其他的例子。

所以一般而言，有必要赋予支付 $p(n)$ 和 $s(n)$ 更宽泛的意义，它并非像在灌溉工程建造例子中那样仅代表成本和收益。我们应允许 $p(n)$ 和 $s(n)$ 是关于 n 的任意函数。这样就不存在关于哪个支付函数比较大的自动的假设，而全部的三种博弈——囚徒困境、懦夫博弈以及安全博弈——以及它们的图形都是值得注意的。事实上，在最一般情况下，$p(n)$和 $s(n)$ 也不一定要是直线，而且可以相交很多次，所以会有几个均衡，并且每一个均衡都可以代表我们之前提到的三种博弈中的一种。[2]当我们在此更一般的意义上进行解释时，我们将提到标记为 P 和 S的行动，但这不一定是指参与或退出，但支付符号会继续使用。所以，当 n 个参与人选择行动 P，$p(n)$ 就成为每个参与人采取行动 P 的支付，而 $s(n)$ 则成为每个参与人采取 S 行动的支付。

12.3　思想简史

很久以前，社会哲学家和经济学家就认识到了集体行动问题。17 世纪英国哲学家霍布斯（Thomas Hobbes）就曾经指出，除非被君主独裁或利维坦（Levi-

athan）统治，否则社会就会在“一切人对一切人的斗争”中瓦解。100 年后，法国哲学家卢梭（Jean-Jacques Rousseau）在其《论人类不平等的起源和基础》（*Discourse on Inequality*）中描述了囚徒困境问题。捕捉一只鹿需要全部猎人合作，将鹿围起来才能杀掉，但任何看见野兔的猎人会觉得放弃合作猎鹿而去追捕野兔是更划算的。卢梭认为此问题是文化的结果，而人类是作为“高尚的野蛮人”在自然界中和谐地生存的。同时，两位苏格兰人对此问题提出了令人振奋的解释：休谟（David Hume）在其著作《人性论》（*A Treatise of Human Nature*）中认为，对未来回报的期望可以维持合作；亚当·斯密在其《国民财富的性质和原因的研究》（*An Inquiry into the Nature and Causes of the Wealth of Nations*）中提出的关于经济的真知灼见，主张纯粹由私人利益带动的物品及服务之生产，将导致社会最优结果。[3]

许多人，尤其是经济学家乃至政治学家，都曾相信这个乐观的解释，认为若结果对团体有利，则团体成员的行动会自动导致这个结果。当 20 世纪 60 年代中期奥尔森（Mancur Olson）出版著作《集体行动的逻辑》（*The Logic of Collective Action*）时，这一信仰受到了强烈冲击。他指出，若要达成集体最优结果，除非每个人所要履行的任务也可给他们带来个人利益——也就是说它必须是一个纳什均衡。然而，他没有把集体行动讲得很清楚。虽然集体行动看起来像囚徒困境，但奥尔森坚持不一定是，当然我们已经知道，它还有可能是懦夫博弈和安全博弈。[4]

另一类主要的集体行动问题——公有资源消耗问题——也在同一时间受到了关注。若牧场和渔业等资源对所有人开放，人人皆可任意使用，则每个人都会尽其所能地使用，因为不用白不用。哈丁（Garrett Hardin）以此为主题写下了名篇《公地的悲剧》（*The Tragedy of the Commons*）。[5]公有资源博弈问题，刚好跟我们前面的例子（部分人参与工程建设，但利益所有人共享）相反。在我们的例子中，每个人都有强烈动机不想参与，指望坐享其成。在公有资源问题中，每个人都有动机想要榨干公有资源，但资源的耗竭会让其他人承担社会代价。

直到最近，许多社会科学家尤其是心理学家在公有资源问题上采取了霍布斯的路线，主张政府强制每个人合作是解决这个问题的唯一方式。而其他人，尤其是经济学家，仍旧认同亚当·斯密的乐观解释，主张对资源正确地界定私有产权，让资源拥有者从中获利，这会诱导资源所有者节制资源使用达到社会最优状态。因为他知道这些资源（草地或鱼类）在未来更不易获得，所以未来价值更高，有必要现在节制使用以便将来获利更多。

而今，各领域的思想家都已认识到集体行动问题有各种形式，而且没有唯一的最优解法。他们也明白社会面临这些问题时并非茫然无助，而是会设计不同的方法来应对。对重复进行的囚徒困境和类似博弈的理论分析宣告了在这方面已经做出许多工作。[6]

12.4 解决集体行动问题

本节简要讨论针对集体行动问题的理论和实际的解决方案，内容主要按照囚

徒困境、懦夫博弈和安全博弈来组织。此三类博弈的共同特点是，需要诱导个人以合作的方式或对集体最优的方式行动，即便个人采取其他行动更符合自己的利益——比如利用他人的合作性行为。[7]

人类表现出了大量的合作行为方式。例如，互惠馈赠行为和防止欺骗的技巧在所有社会及人类历史上屡见不鲜，将其视为本能也是合理的。[8]但人类在从个体成员中诱导集体行为时更加仰仗社会及文化习俗、规范与标准。这些刻意设计的方法是目的在于解决集体行动问题。[9]我们从博弈形式的角度来分析解决这些问题的方法。

12.4.1 分　析

若集体行动问题是安全博弈式的，则解决方案比较简单。此时若预期他人会采取社会最优行动，则自己采取社会最优行动是符合个人利益的。换句话说，社会最优结果是纳什均衡。唯一的问题在于同样的博弈也有另外比较糟糕的纳什均衡。所以，唯一要做的是把最优纳什均衡变成一个聚点（focal point），也就是确保参与人的期望会收敛到聚点。这样的收敛可以来自社会**习俗**（custom）或**惯例**（convention）——也就是自动接纳的行为，因为只要其他人也这么做，就符合每个人的利益。例如，如果某一地区的农夫、牧民、织工（weavers）以及其他生产者想要聚在一起交换他们的器具，他们唯一要做的是确保找到要交易的人。慢慢地，每周在村子 X 日期 Y 开放市集的习俗形成之后，则在市集日聚在一起进行交易就成为每个人的最优选择。[10]

但是 12.2 节的分析已指出，个人的支付组合通常使得集体行动问题，特别是大群体集体行动问题，以囚徒困境的形式出现。所以处理此类问题的方法最受关注就不足为奇了。在第 11 章我们详细介绍了囚徒困境中达成合作的三种方法：重复博弈、惩罚（或奖励）、领导，那时只考虑了双人博弈，但同样的方法加上一些重要的修改和创新就可适用于大群体集体行动问题。

在第 11 章中，我们发现，重复博弈是解决此类问题的最好办法，所以我们对其给予了最大的关注。重复博弈可获得合作性的结果，并且这种合作性的结果就是双人囚徒困境重复博弈的均衡。在这一均衡之下，所有参与人都持有这样的预期，即欺骗将导致合作的瓦解。更一般的情况是，维持合作真正需要的是每个参与人都有这样的预期：欺骗的收益是暂时的，它很快就会被比合作所能获得的收益更低的收益替代。参与人会相信，长期来看，欺骗很快会被遏制，接踵而来的惩罚（未来支付的减少）是迅速、准确、残酷的。

从达成合作这一意义上来看，集体博弈比双人博弈更有优势。同样的两个成员也许没有机会一直频繁互动，但是他们很可能自始至终都与某个人在进行互动。因此，B 也许有欺骗 A 的念头，但出于对将来遭遇其他人（比如 C 和 D）惩罚的担心，这一念头就被打消了。伯拉（Yogi Berra）有一句名言："记住参加别人的葬礼，否则没人参加你的葬礼"，正好描述了一种双方互动不能重复且惩罚必须由第三方实施的极端情形。

但在重复博弈中维持良好行为方面，群体较之直接的二人互动具有补偿劣势。维持合作所要求的对欺骗行为的侦察速度和确定性以及惩罚力度都随成员数量增加而增加。

首先，侦察欺骗行为其实并非易事。在很多现实情形中，参与人的支付并不完全取决于参与人的行动，还会受到某些随机波动因素的影响。即便是两个参与人的情况，若一方得到较少的支付，他也并不能因此确信对方在欺骗他；这也可能是随机冲击的恶果。若人数增加，则会出现另外的问题：如果有人欺骗，他是谁？惩罚一个无法确认其罪责的人不仅在道德上是为人们所排斥的，而且也是不利于生产的——对合作的激励将变得钝化，若合作行为对错误的惩罚比较敏感的话。

其次，在存在大量参与人的情况下，即便欺骗被侦察到并确认出了欺骗者，这一消息必须非常迅速和准确地传达给群体其他成员。因此，群体必须很小，或者有非常良好的沟通或传播网络。而且，成员也不能错误地指控他人。

最后，即使欺骗被侦察到并且该消息可以迅速传遍整个群体，还必须安排对欺骗者的惩罚。而第三方实施惩罚常常需要付出个人的成本——比如，若让 C 去惩罚先前欺骗过 A 的 B，则 C 和 B 之间有利可图的生意就只好作罢，从而实施惩罚本身就是一个集体行动博弈，一样会承受卸责（shirk）之痛；也就是说，没有人去参与实施惩罚。一个社会可以构建一个对卸责者进行处罚的第二轮机制（second-round system），但是那有可能产生新的集体行动问题。然而，人类似乎获得了一种本能，即某些人可以从惩罚骗子中得到快乐，即便他们并不是特定欺骗行为的受害人。[11]

现在我们来讨论社会可用以对背叛者（卸责者）进行惩罚的一些具体方法。一种方法是由团体中其他成员施加**制裁**（sanctions），制裁的方式通常是取消未来参赛的资格。社会还可以制定一系列行为**规范**（norms）改变个人支付诱导合作。这些规范可以以羞耻感、罪恶感或群体中其他成员对背叛者的厌恶之情等形式增加背叛的成本从而改变个人的支付。社会可通过教育或文化来建立规范。规范跟习俗不同之处在于，个人遵守规范不是因为他知道其他人也遵守，而是若违反则有实质的额外代价。规范跟制裁也不同，若违反规范，其他人无须采取明显的行动惩罚你，额外的代价会内化在个人支付中。[12]

社会可以惩罚错误行动，同样也可以奖励正确行动。外界可向该博弈提供金钱等奖励，或者改变参与人的支付，使他们乐于做正确的事（遵守规范）。这两种方式可以一块儿使用。例如，在英国，由于既定的社会规范，人们尊敬贵族和骑士，因此，政府经常授予慈善家贵族称号或骑士头衔以表彰其对社会的贡献。这就是一种外部奖励。

社会大众遵守规范则会使规范更有约束力，人们常常违反规范则会使规范丧失其效力。在福利国家出现之前，当经济处于困难时期时，人们只能依赖其家庭、朋友或社会团体的帮助。这种社会救助一旦成为一种规则就会让人产生惰性，成为依赖他人生活的搭便车者。而当政府发放失业补助时，这种社会救济的功能就会减弱。欧洲在 20 世纪 80 年代末 90 年代初失业率急剧上升，大部分失业人口依赖政府补助生活，使社会救助规则失去了作用。[13]

不同社会或文化团体会形成不同的习俗及规范来达到相同的目的。在一般层面上，每种文化都有自己的一套礼节——跟陌生人打招呼的方式、表达食物可口的用语，等等。当不同文化背景的两个人碰到一起，误会就可能因此产生。更重要的是，每个公司或机构各有其处理问题的方式，它们中的文化与规范的差异微妙且难以达成一致，正是这些“公司文化”冲突导致很多兼并失败。

要成功解决集体行动问题，明显取决于对欺骗和欺骗者的成功的察觉和惩罚。一般而言，小群体中成员与成员间的信息获取较便利，所以察觉欺骗比较容易，惩罚欺骗者也比较容易，所以我们常常看到小社区成功合作的例子，而这在大城市则很难想象。

接下来，考虑集体行动博弈的懦夫博弈形式。采取何种解决方法取决于最大社会总支付是在每个人都参与时（即 12.1 节中的懦夫博弈版本 1）取得，还是在一些人参与其他人退出时（懦夫博弈版本 2）取得。在懦夫博弈版本 1 中，每个人都想退出，这一问题跟在囚徒困境中维持合作类似，所以解决方法也一样；但懦夫博弈版本 2 却不同，既简单又困难，一旦参与者和退出者的角色被定位，他们就没有改变的动机：如果另一人被定位为“径直向前”，则“转向”会让你有更高的支付，反之亦然。因此，若习俗创造了大家对均衡的预期，则该均衡无须进一步的社会干预如制裁就可以维持。然而均衡中退出者的支付比参与者高，这种不公平现象会造成博弈本身的问题。如果僵持与冲突很严重，会威胁整个社会结构。通常重复博弈能解决这些问题，请参与者和退出者轮流互换，平均起来支付就会相等。

有时，囚徒困境博弈版本 2 或懦夫博弈中的支付差异问题不是通过平均支付来解决的，而是以**压迫**（oppression）或**强制**（coercion）的方式让社会中劣势群体接受低支付，并允许优势群体享受高支付。在历史上，处理动物尸体的工作就是强迫由社会某些阶层来完成的；对少数民族以及妇女的不平等待遇也是生动的例子。一旦建立了这样的体制，受压迫群体中没有任何个体可以改变这种状况；受压迫者必须团结起来采取行动改变整个机制，这本身又是另一个集体行动问题。

最后，考虑如何用领导来解决集体行动问题。在第 11 章我们指出，若参与人的“规格”（size）相差悬殊，则囚徒困境会自动消失，因为对于大规格的参与人来说，允许小规格参与人欺骗并继续合作符合其利益。这里，我们还应承认另一种“大”的可能，即“大爱”之人。群体中不同人有不同的偏好，有些人乐于将自己的力量贡献给社会。如果社会上有足够多这样的人，则集体行动问题就会消失。大部分学校、教堂及医院等对社会有重要意义的事业需要这样博爱的人。而这一解决办法同样在小群体中比较有效，因为慈善家会比较容易看到施惠的果实，这样才能鼓励他们继续行善。

12.4.2 应　用

奥斯特罗姆（Elinor Ostrom）在其著作《公共事务的治理之道》（*Governing*

the Commons）中，介绍了几个解决公共资源问题的例子。大部分例子都是采用因地制宜的方法来建立侦察和惩罚机制。例如，土耳其海岸的一个捕鱼社区，其成员轮流在不同渔区捕鱼。某日，某人被分派到了多鱼区，很自然地当有人去侵犯他的地盘时就会被他察觉到并举报。中世纪的英国，通过赋予私人详尽而明晰的权利来避免包括牧场在内的公共资源被过度利用。在某种意义上，这种做法实际上就是将资源分配给个人。

奥斯特罗姆这本书的最大特色就是例子的多样性。她研究的关于公有资源利用的囚徒困境中，一些通过私营化得到了解决，另外一些则因外力或政府的介入而得到了解决。在某些情下，囚徒困境根本无法解决，所有参与人都受困于互相推诿的僵局之中。虽然她举的例子多种多样，奥斯特罗姆还是发现了几个使解决集体行动囚徒困境变得容易的共同特征：（1）要有明确且稳定的潜在参与人；（2）合作的利益要很大，这样监视及强制合作执行的代价才会值得；（3）团体中的成员要能够相互沟通。最后一个特点包含了几个方面：首先，它让规范变得清晰——每个人都知道什么行为是应该的，什么样的欺骗不会被容忍，什么样的制裁会施加到欺骗者头上；其次，它让大家知道侦察机制的效力，这样才能建立互信并消除每个参与人可能的疑虑，即我自己是遵守规则了，可其他人违反规则却逃脱了惩罚。最后，它让团体能监视目前安排的有效性，并根据需要进行改进。这些条件看起来跟我们在第 11 章从对囚徒困境的理论分析（第 2 节和第 3 节）以及对阿克塞尔罗德的竞赛的观察中（第 5 节）所得到的条件一样。

奥斯特罗姆对捕鱼社区的分析也阐明了当集体最优需要不同人做不同的事，而一部分人的支付比其他人高时，可以做些什么。在重复博弈中，轮流交换有利地位，长期下来就是公平的。我们会在第 5 节继续分析公平问题的其他例子以及其他解决方法。

奥斯特罗姆发现合作的外部强制实施者可能难以侦察到欺骗或实施足够明确迅速的惩罚。因而，依靠集权或政府政策来解决集体行动问题的惯常做法常常是错误的。一个例子是来自 19 世纪晚期俄罗斯的乡村社区或“公社”。这些公社正是试图依靠这种方式来解决大量集体行动问题，比如灌溉、庄稼轮植、森林和牧场的管理、道路和桥梁建设与维护，等等。“村社……不是公共协调的港口，农夫的集体行动最符合其个人利益是简单明了的。”20 世纪早期沙皇政府的改革者和 20 年代苏维埃革命家似乎都失败了，部分原因固然是农民头脑里坚持着旧的体制而拒绝一切新的东西，但也是因为改革者未能认识到解决集体行动问题的流行做法的作用，因而未能找出有效的方法去解决问题。[14]

格雷夫（Greif）对中世纪地中海沿岸国家的两个贸易群体进行了比较，揭示了小群体与大群体的差异。马格里布（Maghribis）是依赖于大家族和社会关系的犹太商人的社会组织。如果这个群体的某个成员欺骗了其他成员，受害人就会写信通知其他所有成员。一旦罪责被查实，则群体中就再也没人跟他做生意。这一体制在贸易规模较小的时候运行良好。但是，随着贸易在地中海沿岸扩张，群体就无法继续获得足够亲密和可靠的内部人到存在新的贸易机会的国家去做生意。相反，热那亚商人建立了更正式的法律体制。交易合同必须在热那亚中央当局备案。任何遭遇欺骗或违约的受害人必须向当局申诉，当局会展开调查并对欺

骗者课以适当的罚金。这一体制及其全部的侦察困难，随着贸易扩张而轻易地扩张。[15]随着经济的成长和世界贸易扩张，我们可发现一些类似的转变，比如从紧密联系的群体向更松散的贸易关系转变，以及从基于重复互动的执行向基于官方法律的执行转变。

小群体较能成功地解决集体行动问题的思想，是奥尔森的《集体行动的逻辑》一书的主要议题，并且导致了政治学中的重要发现。在民主社会中，选民具有同等的政治权利，而且主流民意应当取得优势。但是，我们看到太多的例子并非如此。政策对某些群体会产生有利的影响，而对其他群体则会带来不利的影响。要让自己偏好的政策被采用，该群体必须游说、开记者会、营造声势等，而要这样做，该群体就必须解决集体行动问题，因为群体中每个成员可能都希望自己袖手旁观而享受他人努力的成果。如果小群体能解决这些问题，则政策就会偏向小群体一方，即使其他群体成员更多，且遭受的损失超过小群体所得的利益。

最具戏剧性的例子来自贸易政策领域。一国的进口限制会有助于本国产品与进口产品竞争，但这对消费该产品的国内消费者不利，因为进口限制会让产品价格更高。国内的生产者很少，而消费者却几乎占了所有的人口，消费者福利的损失总额远比生产者因此获得的福利总额要大。无论是从经济上考虑利益的得失还是从政治上考虑选民的多寡，进口限制都应该被取消，但我们看到的并不是这个结果。生产者人数少且关系密切，比较能促成政治行动；消费者人数多且分散，所以不易有集体行动。

六十多年前，美国政治家谢茨施奈德（E. E. Schattschneider）提供了大量文件并就高压政治是如何影响贸易政策的这一主题发起了广泛的讨论。他认为“群体的组织能力对它的活动有很重大的影响”，但并没有发展出任何理论来确定群体组织能力。[16]奥尔森和其他人的分析提升了我们对这一问题的理解，但是压力政治胜过经济的事实仍存在于当前的贸易政策中。例如在 19 世纪 80 年代晚期的美国，糖业政策让 2.4 亿美国人每人每年多付出 11.50 美元，总共是 27.5 亿美元，而使 1 万名甜菜农夫每人增加了 5 万美元的收入，以及 1 000 座甘蔗农场每座增加了 50 万美元的收入，总共是 10 亿美元，美国经济的净损失是 17.5 亿美元。[17]每个消费者的损失为 11.50 美元，这对于他们的收入而言是微不足道的，许多消费者甚至不能察觉到。

如果说上述关于集体行动问题的理论和实践综述看起来太分散，并且缺乏一个简洁的总结性陈述，那是因为问题本身就相当分散，并且解决方案依赖于每个问题的具体情况。我们可以提供的一般经验是，让参与人充分意识到利用所处情况的本地信息、其近距离监督其他成员的优势、用不同的持续关系对骗子成员强制实施制裁的能力等设计解决问题办法的重要性。

最后，提醒一句，你可能会从本章关于集体行动问题的讨论中得出这样的结论：个人自由会导致不好的结果，而社会规范及制裁可以改良这一结果。然而，请记住，社会面临的问题不仅是集体行动问题，有些问题恰恰是因为发挥了个体能动性才得以解决，而并非是由于集体的努力。社会常常因受困于规范和习俗而扼杀创新，变得落后和专制。集体行动可能会变成集体不行动，而创新往往是经济增长的关键。[18]

12.5 溢出或外部性

现在我们来比较集体行动博弈中的纳什均衡和社会最优结果。我们试图回答两个问题：为使社会结果达到最优，有多少人该选某个策略，多少人该选其他的策略？而在纳什均衡中，选择各个策略的人数有何不同？

请记住，当总人数 N 中有 n 人选择行动 P，$(N-n)$ 人选择 S，则社会总支付 $T(n)$ 由下式给出：

$$T(n)=np(n)+(N-n)s(n)$$

假如刚开始有 n 个人选 P，而有一个人从 S 转到 P，则选 P 的人会增加到 $(n+1)$，选 S 的人会减少到 $(N-n-1)$，所以社会总支付变成

$$T(n+1)=(n+1)p(n+1)+[N-(n+1)]s(n+1)$$

$T(n)$ 和 $T(n+1)$ 之间的差异就是社会总支付的增加量：

$$\begin{aligned}T(n+1)-T(n)&=(n+1)p(n+1)+[N-(n+1)]s(n+1)\\&\quad-np(n)+(N-n)s(n)\\&=[p(n+1)-s(n)]+n[p(n+1)-p(n)]\\&\quad+[N-(n+1)][s(n+1)-s(n)]\end{aligned}\qquad(12.1)$$

以上是重新整理过后得到的结果。

方程（12.1）有一个很实用的解释，它把一个人从 S 转到 P 的影响做了各种不同的分类。既然一个人相对于整个团体来说是渺小的，其策略改变的影响（或边际影响）也就很小。该方程显示了整个社会支付的边际改变——简称为**边际社会获益**（marginal social gain）——如何分成不同子群体的支付的边际改变。

方程（12.1）中第一项，也就是$[p(n+1)-s(n)]$，是转换策略者的支付的改变；我们称之为**边际私人获益**（marginal private gain），因为它是每个人"私人"持有的。在 12.2 节我们曾见到过类似的表达式。个人选择 S 的支付 $s(n)$ 与选择 P 的支付 $p(n+1)$ 之差即是驱使个人转变的原因，所有的这类个人选择就决定了纳什均衡。

另外，若某人从 S 转到 P，就有 $(n+1)$ 人选择 P；参与者从 n 个增加到 $(n+1)$个，这会改变社会中所有人的支付，不论他们采取 P 还是 S。当一个人的行动影响到其他人，则称其为**溢出效应**（spillover effect）或**外部效应**（external effect），或称**外部性**（externality）；由于我们考虑的是微小变动，我们实际上应称之为边际溢出效应。例如，某通勤驾驶者从路线 A 转到路线 B，会减轻路线 A 的拥挤程度但会加重路线 B 的拥挤程度，也就是减少路线 A 上驾驶者的驾驶时间（增加支付）但增加路线 B 上驾驶者的驾驶时间（减少支付）。虽然一位驾驶者改变路线的效应很小，但如果有很多驾驶者改变路线，整体效应就会很大。

方程（12.1）中第二项和第三项是该个人的改变对其他人的影响。对于其他

n个选择 P 的人，只要一个人转到 P，每人的支付改变量就是$[p(n+1)-p(n)]$，这就是方程（12.1）的第二项；另外还有其他 $N-(n+1)$ 个人仍然选择 S，每个人的支付改变量就是$[s(n+1)-s(n)]$，这就是方程（12.1）的第三项。

因而，我们可以重写方程（12.1），无论对从 S 到 P 或从 P 到 S 的转变都可适用：

边际社会获益＝边际私人获益＋边际溢出效应

例如，若一个人从 S 转向 P，则我们有

边际社会获益$=T(n+1)-T(n)$
边际私人获益$=p(n+1)-s(n)$，以及
边际溢出效应$=n[p(n+1)-p(n)]+[N-(n+1)][s(n+1)-s(n)]$

举个具体例子来解释这些效应。假设你是 8 000 名每日在郊区和市区之间往返的通勤者之一，你可以选择走快线道路（行动 P）或地方道路（行动 S）。走地方道路需要耗费 45 分钟，不管有多少车辆在此道路上行驶。快线道路若无塞车只需 15 分钟。但是每个人选择快线道路都会增加其他每个人的行车时间大约 0.005 分钟（大约 1/4 秒）。

用省下来的时间来衡量支付，即比 1 小时节约多少分钟。这样，走地方道路的支付 $s(n)$ 是一个常数 $s(n)=60-45=15$，无论 n 为何值。但是走快线道路的支付$p(n)$将取决于 n，特别地，若 $n=0$ 则 $p(n)=60-15=45$，但是每增加一个通勤者 $p(n)$ 就降低 5/1 000（或 1/200）。因此，$p(n)=45-0.005n$。

假设最初有 4 000 辆车在快线道路上行驶。这么多的车使得通勤耗时为 $15+4\,000\times0.005=15+20=35$ 分钟，每人得到的支付 $p(n)=25$[即 $60-35$ 或 $p(4\,000)$]。如果你从地方道路转向快线道路，则 4 001 名快线道路行驶者（包括你）将耗时 $35+1/200$ 分钟，即 35.005 分钟；每个快线道路行驶者的支付则变为 $p(n+1)=p(4\,001)=24.995$。这个支付高于走地方道路得到的支付 15。因此你有转向快线道路的私人动机，因为对于你来说 $p(n+1)>s(n)(24.995>15)$。你在转换道路中的边际私人获益 $p(n+1)-s(n)=9.995$ 分钟。但是 4 000 名其他的快线道路行驶者现在因你的决策而多耗费 0.005 分钟；他们每人的支付变化是 $p(4\,001)-p(4\,000)=-0.005$。同样，在地方道路上行驶的通勤者面临的支付变化为 $p(4\,001)-p(4\,000)=0$。对其他司机的累积效应为 $4\,000\times(-0.005)=-20$ 分钟，这是边际溢出效应。总的边际社会获益为 $9.995-20=-10.005$。因此你转变道路的总社会效应是很糟糕的；该例中 $T(n+1)-T(n)<0$，社会支付下降了 10 分钟以上。

12.5.1 一般情况的计算

在详细考察溢出效应弄明白如何做才能达到社会最优结果之前，我们以微积分的语言来重新分析这些概念。如果你不懂微积分，可以跳过不读；但若你懂，

你会发现这比代数方法更简单。

如果群体人数 N 很大，比如成千上万，则每一个人可看做群体中无限小的部分。这使得我们可将 n 看做连续变量。如果 $T(n)$ 是社会总支付，我们就不再用从 n 到 $(n+1)$ 的一个完整单位增加来计算 $T(n)$ 的变化，而是计算 n 改变一个微小量 $\mathrm{d}n$ 后的效应。所以支付改变是 $T'(n)\mathrm{d}n$，$T'(n)$ 是 $T(n)$ 对 n 的导数。社会总支付的表达式为

$$T(n)=np(n)+(N-n)s(n)$$

微分后我们得到

$$\begin{aligned}T'(n)&=p(n)+np'(n)-s(n)+(N-n)s'(n)\\&=[p(n)-s(n)]+np'(n)+(N-n)s'(n)\end{aligned} \quad (12.2)$$

上式跟方程（12.1）是等价的。$T'(n)$ 表示边际社会获益。边际私人获益为 $p(n)-s(n)$，这是从 S 转到 P 的个人支付改变。在方程（12.1）中我们将支付改变表示为 $p(n+1)-s(n)$，而现在是 $p(n)-s(n)$，这是因为微小增加量 $\mathrm{d}n$ 并不会显著改变每个选择 P 者的支付。但所有选择 P 者的支付加总 $np'(n)$ 是很可观的，需要纳入溢出效应计算——这是方程（12.2）中的第二项。$(N-n)$ 个选择 S 者的支付改变也是如此，$(N-n)s'(n)$ 是方程（12.2）的第三项，而最后两项组成了边际溢出效应。

在通勤路线例子中，我们有 $p(n)=45-0.005n$，以及 $s(n)=15$。那么，若使用微积分，转向快线道路的通勤者的边际私人获益仍使用 $p(n)-s(n)=30-0.005n$ 来计算；因为 $p'(n)=-0.005$ 而 $s'(n)=0$，溢出效应为 $n\times(-0.005)+(N-n)\times 0=-0.005n$，当 $n=4\ 000$ 时溢出效应为 -20。这个答案与先前的结果一样，但微积分简化了推导并有助于我们直接发现最优解。

12.5.2 负溢出

在通勤的例子中，边际溢出效应是负的（-20 分钟），我们说这种状况存在负外部性。当一个人的行动会降低其他人的支付时，负外部性就会存在；它把一些额外的代价强加给了社会中的其他人，但行为人却只是受到其自身支付的激励。（记住，他在伤害其他人时产生的罪恶感也已经反映在其支付中。）因此他在进行决策时不会考虑溢出效应。只要从 S 行动转换到 P 行动可以获得正的边际私人获益，他就一定会转换。如果个人决策由边际社会获益决定，则社会的景况可变得更好。在我们的例子中，边际社会获益是负的，但是边际私人获益是正的，所以个人将转换其行动，尽管他若不转化则社会的景况会更好。更一般地，在存在负外部性的情况下，边际社会获益将小于边际私人获益。人们基于成本－收益计算而制定的决策从社会角度来看是错误的；相比于社会所期望的，个体参与人更为经常地选择具有负溢出效应的行动。

我们可用方程（12.1）来计算特定个人与整个社会可从行动转换中受益的精

确条件。请回忆一下，若已经有 n 个人使用快线道路，而另一个人正在考虑从地方道路转换到快线道路上，如果 $p(n+1)>s(n)$，他就可从转换中得到好处，而若 $T(n+1)-T(n)>0$，则总的社会支付就会增加。私人获益在以下条件成立时为正：

$$45-(n+1)\times 0.005>15$$
$$44.995-0.005n>15$$
$$n<200(44.995-15)=5\ 999$$

社会获益为正的条件为

$$45-(n+1)\times 0.005-15-0.005n>0$$
$$29.995-0.01n>0$$
$$n<2\ 999.5$$

所以，如果允许自由选择，通勤者将拥挤到快线道路上，直到接近 6 000 个人为止，但只要超过 3 000 人总社会支付就会下降。所以快线道路上的驾驶者保持在 3 000 人以下对于社会来说是比较有利的。

我们用图 12—4 来表示上述结果。横轴自左至右衡量的是快线道路行驶的通勤者的数量。纵轴衡量的是到快线道路上有 n 个通勤者时，每个通勤者的支付。每个走地方道路者的支付是常数，无论 n 为多少其支付皆为 15，以水平线 $s(n)$ 表示。每个转换到快线道路者的支付以 $p(n+1)$ 表示，n 每增加一单位则该支付下降 0.005。两条线在 $n=5\ 999$ 处相交；即，$p(n+1)=s(n)$ 或在边际私人获益为零时的 n 值处相交。在交点左边的任何 n 值上，在地方道路上的每一个行驶者若转向快线道路都可获得正的收益。随着一些人的转换，快线道路上人数增加——也就是 n 值增加。相反，在交点右边（即 $n>5\ 999$），$s(n)>p(n+1)$；快线道路上（$n+1$）个行驶者会因转到地方道路而获得好处。随着一些人的转换，

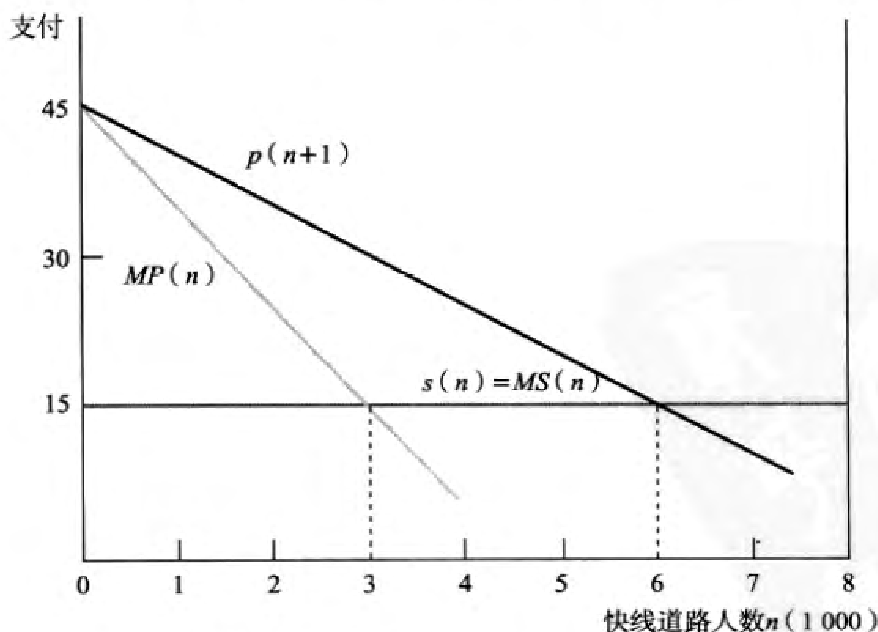


图 12—4　路线选择博弈的均衡和最优结果

快线道路上行驶者数量 n 会减少。在交点左边，会收敛到 5 999 而在交点右边会收敛到 6 000。

如果我们使用微积分方法，我们可把 1 视为非常微小的增加量，所以 $p(n)$ 就取代 $p(n+1)$，而交点就是 6 000 而不是 5 999。你会发现，这实际上几乎没有差别，我们可把 $n=6\,000$ 视为纯粹考虑个人情况下路线选择博弈的纳什均衡。给定自由选择，8 000 人中的6 000人将选择走快线道路，而只有 2 000 人走地方道路。

但我们也可从整个社会的观点来解释博弈结果。当 $T(n+1)>T(n)$ 时，社会就会因快线道路行驶者数量增加而获得好处。当 $T(n+1)<T(n)$ 时，社会就会有损失。为了用图形表示，我们将方程（12.1）分拆成两个部分，一部分是 p 的函数而另一部分是 s 的函数：

$$\begin{aligned}T(n+1)-T(n)&=(n+1)p(n+1)+[N-(n+1)]s(n+1)\\&\quad-np(n)-(N-n)s(n)\\&=\{p(n+1)+n[p(n+1)-p(n)]\}\\&\quad-\{s(n)-[N-(n+1)][s(n+1)-s(n)]\}\end{aligned}$$

上述表达式第一部分是选 P 者的支付效应，包含了转换者的支付 $p(n+1)$ 及溢出效应 $n[p(n+1)-p(n)]$，我们把这称为从 n 到（$n+1$）时，P 子群体的边际社会支付，或简称 $MP(n+1)$。同理，第二部分是 S 子群体的边际社会支付，简称 $MS(n)$。如果 $MP(n+1)>MS(n)$，某人从 S 转到 P 会让社会总支付增加（或某人从 P 转向 S 会让社会总支付减少），如果 $MP(n+1)<MS(n)$，某人从 S 转到 P 会让社会总支付减少（某人从 P 转向 S 会让社会总支付增加）。

结合我们在通勤者例子表示 $p(n+1)$ 及 $s(n)$ 的方法，我们有

$$\begin{aligned}MP(n+1)&=45-(n+1)\times0.005+n\times(-0.005)\\&=44.995-0.01n\end{aligned}$$

而对于任何 n 皆有 $MS(n)=15$。图 12—4 表示了 $MP(n+1)$ 和 $MS(n)$ 的关系。注意 $MS(n)$ 与 $s(n)$ 永远重叠，原因是地方道路永远不会堵塞。但是 $MP(n+1)$ 位于 $p(n+1)$ 曲线之下，原因是存在负的溢出效应，个人转到快线道路时社会获得的好处比转换者私人获得的好处要少一些。

$MP(n+1)$ 和 $MS(n)$ 两条曲线在 $n=2\,999$ 或近似 3 000 处相交。交点左边 $MP(n+1)>MS(n)$，社会因多一个人转换到快线道路而获益；交点右边情况则相反，社会将因多一个人从快线道路转换到地方道路而获益。所以社会最优配置是 3 000 个行驶在快线道路上，3 000 个行驶在地方道路上。

如果你愿意使用微积分，可以把快线道路行驶者的总支付写成 $np(n)=n(45-0.005n)=45n-0.005n^2$，于是取导数可得到 $MP(n+1)=45-0.005\times2n=45-0.01n$。剩下的分析跟以前一样。

社会如何达到行驶人数的最优配置？不同的文化和政治团体使用不同的方法，各有其优劣。社会可简单地将快线道路行驶人数限制在 3 000 人。但问题是如何选择这 3 000 人？可以采取先到先得的办法，但驾驶人员就会互相尽早占据快线道路，而这又会浪费时间。官僚社会可能会设立一些标准让有优先权的人士

或有需要的人士走快速道路，但这样一来人们便会采取一些代价不菲的行动以达到标准。在政治化的社会，标准的选择可能会被利益团体所左右。在腐败的社会，官员及贿赂官员的人可以走快线道路。在平等主义社会，可能由抽签决定，或是每月轮流。一个例子是依车牌的最后一位号码决定哪些天可以走快线道路，但这似乎也不公平，因为有钱人可以有好几辆车，这样他就可以每天选择不同的车牌号码走快线。

许多经济学家偏爱另外一种更为开放的收费制度。假设走快线道路的人要缴税 t，以时间衡量，则走快线道路者的个人获益变成 $p(n)-t$，n 是纳什均衡数量[在此我们忽略 $p(n+1)$ 和 $p(n)$ 的微小差别，因为 N 很大] 并由 $p(n)-t=s(n)$决定。我们知道社会最优的 n 值是 3 000，运用表达式 $p(n)=45-0.005n$ 和 $s(n)=15$，将 n 取 3 000 代入，可得到驾驶人员对快线道路和地方道路保持中立的条件：$45-15-t=15$，或 $t=15$。如果以最低工资每小时 5 美元衡量，15 分钟就值 1.25 美元。只要对快线道路行驶者征收 1.25 美元的税收或过路费，就可以将快线道路行使人数控制在社会最优数量。

注意，如果已经有 3 000 个人行驶在快线道路上，增加一个人将使每个人耗时增加 0.005 分钟，总共是 15 分钟。这正好就是每位上快线道路的驾驶员需要缴纳的过路费。换句话说，每位驾驶员都要承担自己给社会造成负溢出效应的代价，所以这会诱导他采取社会最优行动。经济学家把这称作**外部性的内部化**(internalize the externality)。损害别人者，要承担损害的责任，这一思想很有吸引力。但是过路费不是用于补偿其他人的，如果这样的话，每个走快线道路的人都希望收到其他人的过路费，收费的初衷就变质了，所以过路费要缴入国库，用于有益于社会的用途。

其他一些偏爱用市场方式解决问题的经济学家则认为，如果道路是私人的，则他牟利的动机会令他对使用者征收费用，且会刚好让使用者数量达到社会最优标准。如果他对使用者收费 t，则使用者数量由 $p(n)-t=s(n)$决定，他的利益是 $tn=n[p(n)-s(n)]$，并且他试图最大化这一利益。在例子中，收益是 $n[45-0.005n-15]=n[30-0.005n]=30n-0.005n^2$；很容易知道 $n=3\,000$ 时收益达到最大。但是，这种情况下，收益全部落进了私人的腰包；绝大多数人都认为这是一个不好的结果。

12.5.3 正溢出

关于正溢出或正外部性的诸多问题，可以视为负溢出的镜像（mirror）来了解。当个人的行动有正溢出时，他的边际私人获益就会比边际社会获益小，所以该行动不会被充分采取，而且他在纳什均衡状态下的收益也是不充足的。所以需要想办法增加激励——提供跟溢出利益相等的奖赏，就能达到社会最优结果。

的确，正溢出和负溢出的区别只是字面上的，是正溢出还是负溢出取决于哪个行动叫做 P，哪个行动叫做 S。在通勤者的例子中，若我们把地方道路叫做 P 而把快线道路叫做 S，则通勤者从 S 转到 P 会减少其他走快线道路者的时间，所

以具有正溢出。另一个例子，考虑针对某传染病的疫苗，每个注射疫苗的人会降低得病的风险（边际私人获益），也会降低传染给其他人的风险（溢出）。如果没注射疫苗是 S，则注射疫苗就有正溢出效应；如果没注射疫苗是 P，则从 S 转到 P 有负溢出效应。这些都暗示了可以设计策略让个人行动符合社会最优结果，比如奖励注射疫苗的人，惩罚没注射疫苗的人。

但是，具有正溢出的行动跟具有负溢出的行动相比，有一个非常重要的特点，**即正反馈**（positive feedback）。假设你选择 P 的溢出效应会增加其他人选择 P 的支付，则你的选择增加了选择 P 的吸引力，引诱其他人也选择 P，而其他人选择 P 也引诱你选择 P，互相反馈的结果是每个人都选择了 P。相反地，如果很少人选择 P，没有吸引力的结果使其他人也放弃 P，最后造成所有人都选择了 S。换句话说，正反馈会产生多重纳什均衡。我们现在用一个实际例子来加以说明。

假设你要买一台电脑，你需要在 Windows 操作系统以及 Unix 操作系统之间进行选择。这个博弈有其必要的支付结构。若 Unix 用户增加，购买 Unix 系统的电脑就更好——该系统的漏洞会更少，因为大量的用户会察觉到漏洞的存在而加以弥补，也可以获取到更多的应用软件，而且出现问题时也可找到更多的专家来帮助解决。同样，若 Windows 用户越多，则基于 Windows 系统的电脑就越有吸引力。当然，许多电脑迷可能认为 Unix 系统更优越，不过这不是我们所关心的问题，我们关心的是前述情况下究竟会发生什么，个人选择是否会导致社会最优结果。

用类似图 12—1 到图 12—3 的图（图 12—5）可以表示出个人电脑购买者使用两个策略的支付，两个策略分别是购买 Unix 电脑和购买 Windows 电脑。如图 12—5 所示，购买 Unix 电脑的支付随 Unix 用户数量增加而增加，而购买 Windows 电脑的支付随 Unix 用户数量下降（Windows 用户数量增加）而增加。正如我们已经解释过的，图中每个人皆为 Unix 用户时其支付（标记在点 U）比每个

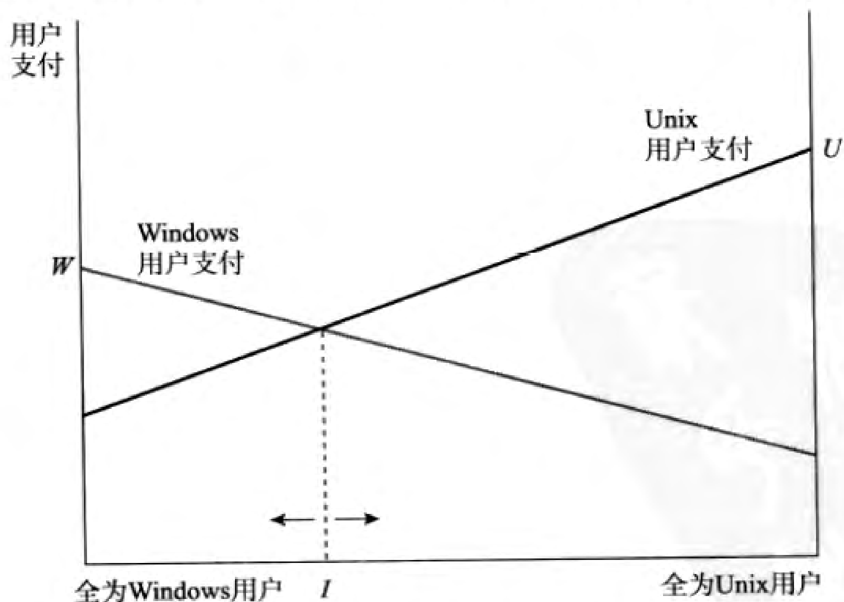


图 12—5　操作系统选择博弈的支付

人皆为 Windows 用户时其支付（标记在点 W）要高。

若当前的人口中只有少数 Unix 用户，这种情况在图中表示为两条支付线交点对应的 I 点之左侧，每个用户都会发现选择 Windows 操作系统是更好的。若 Unix 用户比较多，这种情况对应于 I 点之右侧，那么每个人选择 Unix 操作系统是更好的。因而在当前情况下，只有当 Unix 用户数量为 I 时，才会达成既有 Unix 用户又有 Windows 用户的均衡。只有在此时才没有人有动机转换其角色。但这种状态并不稳定。可以设想，刚好有一个人由于偶然做出了不同的选择，若他转向 Windows，他的选择就将使用 Windows 操作系统的人数推向了 I 的左边，而这又会刺激其他人也转向 Windows 操作系统。如果他转向 Unix 操作系统，使用 Unix 的人数将被推向 I 的右边，这会刺激人们转向 Unix 操作系统。这些转换的累积效应最终将把整个社会推向全部使用 Unix 操作系统或全部使用 Windows 操作系统的结果；这是两个稳定的均衡。[19]

但是哪个稳定均衡会产生？这取决于博弈的起点。如果你观察今日电脑之使用者，你会发现大部分人都使用 Windows 操作系统。因此，由于 Unix 用户太少了（或者太多的个人电脑用户），整个世界正向着全部使用 Windows 操作系统的均衡演进。仅仅由于机缘巧合，学校、企业以及个人用户都已经被**锁定**（lock in）在这个特定的均衡。若每人皆使用 Unix 操作系统的确可以给社会带来更大的收益，则我们应该使用的就应是 Unix 操作系统而不是 Windows 操作系统；但不幸的是，没有一个人想改变这种状况。只有协调行动才能扭转到全部使用 Unix 操作系统的均衡，也就是超过 I 的人使用 Unix 操作系统，才会让其他人也理性选择 Unix 操作系统。

还有类似的由不同群体进行的习惯选择的例子。最有名的案例是那些做了错误选择的案例。有人倡导蒸汽机比汽油引擎的效率高——它当然会更干净。Dvorak 键盘的支持者认为他们的键盘比 QWERTY 键盘好。许多工程师认为 Betamax 视频录像机比 VHS 视频录像机要好。这些情况中，人们一时的错误念头或天才的广告商会决定最后均衡结果，但这一结果从社会整体角度来看可能是不好的。不过也有些情况的均衡结果是好的，例如很少有人会觉得红绿灯的颜色有何不妥。[20]

正反馈以及锁定的思想在宏观经济学中非常重要。当国民收入提高时，需求就越大，而生产利润也就越高。反过来，工厂生产越多，就需要雇用更多的员工，国民收入就越高。这样的正反馈创造了多重均衡，高所得、高生产是社会的较佳结果，但个人决定会将社会锁定在低收入、低生产的均衡。在本章的习题 9 中，我们要求你将上述思想用图表示出来。较佳的均衡可通过公开宣称“我们唯一害怕的事情就是‘害怕’本身”来使之成为聚点，但是政府也可适当地扩大政府支出规模增加需求从而使社会趋向较佳的均衡。换句话说，总需求不足会导致失业——英国经济学家凯恩斯（John Maynard Keynes）曾在 1936 年的巨著《就业、利息和货币通论》中用经济学理论中的供给需求模型加以讨论过；从博弈论的角度来看，可将它视为解决集体行动问题失败的结果。[21]

12.6 “救命!”：有混合策略的懦夫博弈

在前面的懦夫博弈式集体行动问题中，我们只讨论了纯策略均衡。但是我们从 8.1.2 节知道这类博弈也存在混合策略均衡。在集体行动问题中，每个人都在想：“等其他人参与建造，然后我就可以偷懒；但是他们可能不会参与建造，这样我就必须参与建造。”混合策略巧妙地抓住了犹豫的心态，下面有一个令人印象深刻的例子。

在 1964 年的纽约市（皇后区的 Kew 花园），一个叫吉诺维斯（Kitty Genovese）的妇女被歹徒杀害。残忍的袭击持续了半个小时，她一直在尖叫，很多人也听到了她的尖叫声，超过 30 人在命案现场，但没有人帮助她，也没有人报警。

这件凶杀案引起了轰动，并有几种理论对其进行了解释。新闻界以及大部分公众都认为纽约人——或大城市居民，或美国人，乃至所有人——对他们的同胞冷漠无情。

然而，稍微观察一下就会让你相信，其实大家还是很关心自己的同胞的，即使是陌生人。社会学家对情况做了不同的解释，称为多元无知（pluralistic ignorance）。因为没有人知道发生了什么事，是否要帮助而又要帮助多少，他们互相看着对方寻找线索，并尝试解读其他人的行为。如果没有人去帮助，他们就解读成这位妇女不需要帮助，所以他们就不帮助她。

这似乎有道理，但却不能完全解释吉诺维斯这样的命案。尖叫的妇女需要帮助，这是强烈的合理假设。旁观者究竟在想什么——在拍电影？如果是的话，那灯光呢？摄像机呢？其他演员呢？

有种解释比较好：每个旁观者会因吉诺维斯的受害而难过，也会因她获救而高兴，但帮助吉诺维斯要付出代价，比如报警时要出示身份证明，还要充当证人等等，所以我们看到大家都宁愿等别人报警解救吉诺维斯，这样自己就能不付出任何代价，而得到高兴的收益。

社会心理学家对这种希望不付代价获益的思想有不同的解释。他们称之为**责任分散**（diffusion of responsibility），也就是大家都同意帮助是必要的，但他们彼此不沟通，无法协调出由谁来提供帮助，每个人都认为帮助是他人的责任，而且团体越大，每个人就越以为有其他人会出来帮助，从而自己就可以省却麻烦。

社会心理学家做了一些实验来检验这一假说。他们在不同规模的人群中设计出某些需要他人帮助的情形，结果发现人群越大，就越得不到帮助。

责任分散的思想似乎可以解释这种情况，但也不尽然。它说人数越多则大家越不可能帮助。不过虽然人多，但只需要一个人报警即可。所以要让帮助的可能性降低，必须让每个人帮助的概率加速减少，才能盖过人数增加的效应。而要知道是不是如此，我们必须用博弈论来分析。[22]

我们只从责任分散的观点来看，也就是旁观者的行动并没有什么协调，并忽

略其他信息的干扰，再假设每个人都相信帮助是必须且值得的。

假设有 N 个人，每个人皆可因为帮助而获得收益 B。只需要一个人帮助，多则无益。任何实施帮助的人都需要承担代价 C，我们假设 $B>C$，所以即使没有人愿意帮助，自己去帮助也是值得的，这样帮助行动才能显得合理。

某个采取帮助行动的人会得到收益 B，付出代价 C，净支付是 $(B-C)$；如果是其他人行动，则他会得到更高的支付 B。所以每个人都希望其他人行动，这样他就可以白捡便宜得到更高的支付。但若所有人都这样想，均衡结果会是什么？

如果 $N=1$，就不算博弈，而只是一个单人决策问题。这个人采取行动会得到 $B-C>0$，不采取行动就什么也没有，所以他会行动。

如果 $N>1$，就变成了有多个均衡的策略互动博弈。我们可以先排除一些可能性。$N>1$ 时，不可能存在全部人都行动的纯策略均衡，因为一定有人变成旁观者等待不付出代价而获益。同样，也不可能有全部人都不行动的纯策略均衡，因为大家知道其他人都不行动后（请记住纳什均衡假设下，每个参与人都把其他人的策略视为给定），一定会有人行动以获得支付。

所以只剩下一个人采取行动的纳什均衡，事实上有 N 个这样的均衡，因为每个人都可以是那唯一的采取行动者。但个人在孤立地做出个人决定时，无法协调指派谁采取行动。即便有这种协调，大家也可以相互推卸责任，达不成一致意见，或是一致意见尚未达成吉诺维斯就已遇害了。所以我们必须考察对称的均衡，其中每个人有相同的策略。

我们已经明白不可能存在 N 个人全部都采取相同的纯策略的均衡，因此我们应看看能否寻找到每个人都采取相同混合策略的均衡。事实上，在这里很适合用混合策略，每个人都是独立的，而且都在猜想其他人怎么行动：我是否应该报警？……不过别人可能会报警……但若他们不报警呢？……这一连串的想法在每个人的脑海中旋转，而他们的行动是最后闪过的念头，但我们无法预测是什么，所以用混合策略代表他们的随机决定。

假设 P 是任意某一人不行动的概率，如果某人使用混合策略，行动与不行动这两个纯策略对于他来说必须是无差异的。行动会让他得到（$B-C$）；他不行动且其他（$N-1$）个人也不行动则得到 0，其他（$N-1$）个人至少有一个人行动得到 B。既然每个人不行动的概率都是 P，而决定是独立做出的，所以其他 $(N-1)$ 个人都不行动的概率为 P^{N-1}，而其中至少有一个人行动的概率是（$1-P^{N-1}$），因此某人不行动的期望支付是

$$0\times P^{N-1}+B(1-P^{N-1})=B(1-P^{N-1}).$$

而他对行动与不行动感到无差异的条件是

$$B-C=B(1-P^{N-1})\text{ 或 }P^{N-1}=C/B\text{ 或 }P=(C/B)^{1/(N-1)}.$$

注意，这一特定个人的无差异条件也适用于决定其他参与人混合其策略的概率。

得到混合均衡概率后，我们现在看看概率如何随 N 的改变而改变。因为 $C/B<1$，所以随着 N 从 2 增加到正无穷，$1/(N-1)$ 就从 1 减到 0，P 就从 C/B

增加到 1。记住，P 是任意某一人不行动的概率，所以任意某一人行动的概率 $(1-P)$ 将从 $(1-C/B)$ 递减到 0。[23]

换句话说，旁观者越多，每个人就越不可能报警。这在直觉上是成立的，而且符合责任分散的思想。但还不能告诉我们是否越多旁观者在场，吉诺维斯就越不可能得到帮助。正如我们之前说过的，只需要有一人出面帮助即可，人越多每个人就越不可能帮助，但这不等于至少有一人帮助的概率也越小。我们需要更多的计算来说明。

由于在纳什均衡中，N 个人的决策是随机且独立的，则没有任何人帮助的概率 Q 是

$$Q=P^{N}=(C/B)^{N/(N-1)}$$

随着 N 从 2 增加到正无穷，$N/(N-1)$ 从 2 减少到 1，从而 Q 也从 $(C/B)^2$ 增加到 C/B。相应地，至少有一人帮助的概率 $(1-Q)$ 会从 $1-(C/B)^2$ 减少到 $(1-C/B)$。[24]

我们用精确的计算证明了假说：群体越大，越不可能提供帮助。但是即使非常多的人在场，得到帮助的概率也不会降低到 0，而是趋于一个正值，也就是 $(B-C)/B$，这取决于个人行动的利益与代价。

现在我们已明白博弈理论分析如何深化了来自社会心理学的思想。责任分散理论部分地得到了我们的结论，即越大的群体中每个人越不可能提供帮助；但若要得到更进一步的结论——越大的群体中越不可能提供帮助——则需要更精确的概率计算，这是根据混合策略和互动均衡的分析而来的。

现在我们要问，吉诺维斯死有余辜吗？根据多元无知和责任分散的理论，越大的城市中，每个人就越冷漠吗？搭便车的游戏会继续上演吗？也许不是。《纽约时报》（*New York Times*）的记者泰尔尼（John Tierney）公开撰文赞扬“都市怪人”的美德[25]，有一些怪人会惩罚没有公德心的人——包括环境污染者、行为粗鲁者等，这些人本质上是社会规范的强制者，虽然社会上已经有这些怪人，但泰尔尼还是呼吁我们应加入怪人的行列。读他的文章可能没有意义，因为社会规范还是在四处被践踏，毕竟世界上并非只有你一人！你今天有没有迫使别人遵守规则呢？换句话说，我们需要社会规范，也需要有些人会因为规范的强制执行而从中获得支付。

12.7 总　　结

多人博弈通常牵涉到集体行动问题。集体行动博弈的一般结构可能表现为囚徒困境、懦夫博弈或安全博弈。此类博弈中，任何形式的博弈的关键问题都是出于个人理性选择而导致达成的纳什均衡不一定是社会最优结果——也就是令所有参与人的支付总和最大的结果。

很早以前，不同领域的学者就认识到并讨论了集体行动问题。刚开始有人认为该问题无法解决，也有人提出了令人振奋的解决方案。19 世纪 60 年代，奥尔森认为，除非个人有动机去执行他被指派的工作，否则不能达到社会最优结果。最近人们则认为，集体行动问题来自各种领域，而且没有单一的最优解决方案。

社会科学分析表明，社会习俗或惯例可以导致合作行为。其他的可能解决方案来自对非合作行为的制裁或者对可接受的行为规范的创造。大部分的理论都同意小群体比大群体更能解决集体行动问题。

在集体行动博弈中，若某个人的行动对其他所有人的支付有影响，我们就说存在溢出或外部性。溢出可以是正的也可以是负的，但都不能导致社会最优结果：当行动产生负溢出时，从社会的角度看该行动会被过度采取；当行动产生正溢出时，则该行动被过少地采取。可以使用让个人为其行为付出代价（或获得好处）的机制来达到社会最优结果。另外，存在正溢出时会产生正反馈，这种情况下，博弈有多重纳什均衡。

大群体博弈中，责任分散会让人袖手旁观，等待别人的行动，白捡便宜获得好处。在需要帮助的时候，群体越大似乎越不可能获得帮助。

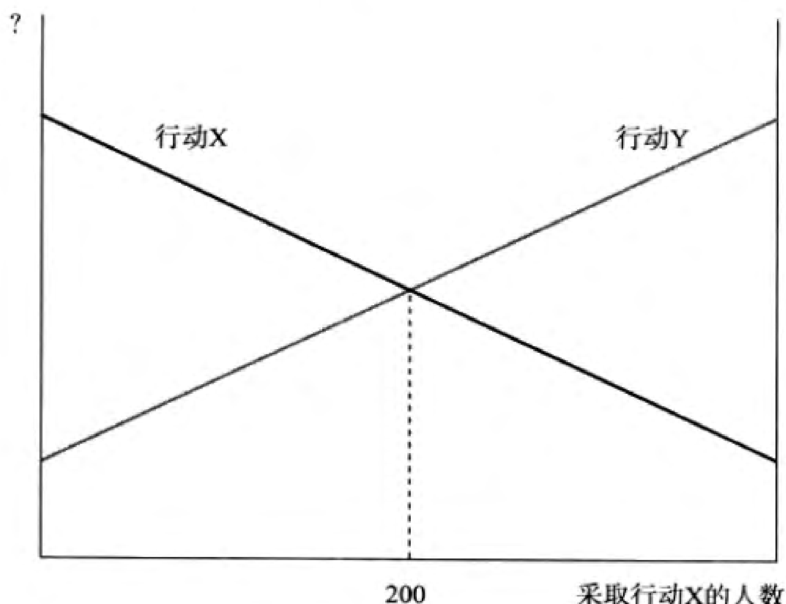
关键术语

抑制（coercion）
集体行动（collective action）
惯例（convention）
习俗（custom）
责任分散（diffusion of responsibility）
外部效应（external effect）
外部性（externality）
搭便车者（free rider）
外部性内部化（internalize the externality）
锁定（lock in）
边际私人获益（marginal private gain）
边际社会获益（marginal social gain）
非排他性利益（nonexcludable benefits）
非竞争性利益（nonrival benefits）
规范（norm）
压迫（oppression）
多元无知（pluralistic ignorance）
正反馈（positive feedback）
纯公共物品（pure public good）
制裁（sanction）
溢出效应（spillover effect）

习　题

1. 假设有 400 人正在行动 X 和行动 Y 之间进行选择。两个行动的相对支付取决于 400 人中多少人选择行动 X 多少人选择行动 Y。支付如下图所示，但是纵轴未做标记，所以你不知道纵轴表示的究竟是两种行动的收益还是成本。
   (a) 若告诉你 200 人选择行动 X 是非稳定的结果。那么，若现在有 100 人选择行动 X，则你估计选择 X 的人数会增加还是减少？为什么？
   (b) 为了与你在（a）中描述的行为保持一致，则纵轴应当标记为行动 X 和行

动 Y 的成本还是收益？解释你的答案。



2. 表 12—5 表示了一般的二人集体行动博弈的支付。在那里我们用代数支付给出了导致囚徒困境的各个不等式。现在，请你找出其他类型博弈相对应的不等式。
   (a) 在何种支付条件下，双人博弈是一个懦夫博弈？导致懦夫博弈版本 1（表 12—3）的进一步条件是什么？
   (b) 何种支付条件下，双人博弈是一个安全博弈？
3. 一个有 30 名学生的班级，布置了有 5 道问题的家庭作业。前面 4 道题普普通通，但第 5 道题是一个互动博弈。题目写道："你可以选择是否回答此题。如果选择回答，你只需写上'我以此方式回答问题 5'。如果你选择不回答问题 5，你的作业成绩将根据前 4 道题的表现打分。如果你选择回答问题 5，则你的得分如下：若全班少于一半的学生回答了问题 5，你从第 5 题得到 10 分；这 10 分将累加在前 4 道题作业的得分上。若超过一半的学生回答了问题 5，你将得到 -10 分，即从其他题目得分中扣掉 10 分。"
   (a) 画图说明"回答问题 5"和"不回答问题 5"两种可能策略与回答问题人数相联系的支付。找出博弈的纳什均衡。
   (b) 若该博弈在大学教室里真正进行，你估计会产生什么结果？为什么？考虑两种情况：（ⅰ）每个学生不进行沟通各自选择，以及（ⅱ）学生各自选择，但选择前允许在班级网站上进行讨论。
4. 某团体有 100 个成员。若他们中 n 人参加了一个公益项目，每个参加者所得支付为 $p(n)=n$，并且 $(100-n)$ 个偷懒者每人所得支付为 $s(n)=4+3n$。
   (a) 这是什么类型的博弈？
   (b) 写出团体总支付的表达式。
   (c) 证明 $n=74$ 时会产生团体最大总支付。
   (d) 要让 74 人参加而剩下 26 人退出，会遇到什么困难？

(e) 团体要怎么克服这些问题?

5. 考虑一个总人口为 100 万的小地区，每个人都可选择在阿尔法和贝塔两个城镇之一居住。对于每个人而言，居住在某个城镇的利益会随村镇人口数增加而增加（因为人越多则公共设施越多），但是当人口超过某一数量，利益就会显著减少（因为拥挤）。如果 x 是跟你同住一城镇的人口比例，则你的支付是：

x，如果 $x<0.4$

$0.6-0.5x$，如果 $0.4<x<1$

(a) 画图表示出上述关系。x 可看做 0 到 1 之间的连续变量。

(b) 均衡可以在两种情况下达到：两个城镇都住人，而他们的支付一样，或是某城镇（例如贝塔村）完全没住人，而另一城镇（阿尔法村）居民的支付会比第一个搬到贝塔村居住者支付高。在图上找出这两个均衡。

(c) 现在考虑一个动态调整过程，因为某城镇居民的支付暂时较高，所以大家逐渐迁移到该城镇。动态的均衡已经在 (b) 中讨论过了，试问，哪个是稳定的均衡，哪个是不稳定的均衡?

6. 从 A 到 B 有两条驾驶路线。一条是高速公路，一条是地方公路。使用高速公路的利益是常数，等于 1.8，不受走高速公路的人数的影响；太多人走地方道路的话会比较拥塞，太少人走地方公路可能会被强盗劫持。假设走地方公路的人数比例是 x，则走地方公路的利益是

$$1+9x-10x^2$$

(a) 画图表示两条路线的支付，以函数 x 表示，x 是 0 到 1 之间的连续变量。

(b) 从 (a) 的图中确认所有可能的均衡交通状态，其中哪些是稳定的，哪些是不稳定的，为什么?

(c) 哪个 x 让全部人口的总支付最大?

7. 假设在某个有 100 人的城市建造一个游乐园，需要大家捐钱，要求每位市民捐 100 美元，捐得越多则游乐园可修得越大，每位市民的支付也就越大。但还是有些人偏就不捐，而且不捐者也会得到好处。假设有 n 个捐献者，n 是从 0 到 100 中的任何数字，每位市民的利益等于 n^2，以金钱为单位来衡量。

(a) 假设刚开始没有人捐，而你是市长，想要每个人都捐钱。你可以游说某些人。请问你要说服多少人捐献，才能让其他所有人自愿捐献?

(b) 找出此博弈的纳什均衡，在均衡中每个人都在犹豫是否捐献。

8. 假设一个班有 100 名同学，正在两种职业之间进行选择：律师或工程师。一个工程师每年可赚到手的钱是 100 000 美元，而不管多少学生选择了这个职业。律师每年赚到的钱，随律师数量增加而增加，到达某个点后，随着竞争的增加收入会下降。具体地说，若有 N 个律师，则每个律师每年可获得 $1\,000(100N-N^2)$ 美元。提供法律服务的年成本（办公室租金、秘书薪水、助理律师薪水、互联网文献服务费用等）为 800 000 美元。所以每个律师每年可赚到手的钱是 $1\,000(100N-N^2-800)$。

(a) 画图表示每个律师赚到手的钱（纵轴表示）和律师的数量（横轴表示），

描绘出一些点，比如 0，10，20，…，90，100 个律师，并用计算机绘图程序（如果你有的话）拟合成一条曲线。

(b) 如果职业选择在无协调的方式下进行，那么可能的均衡结果是什么？

(c) 现在假设全班统一决定多少人从事律师职业以最大化全班同学每年赚到手的钱的总数，则从事律师职业的学生数量是多少？（如果可以，请使用微积分，视 N 为连续变量。或者你可以使用图形方法或者用电子表格计算。）

9. 把 12.5.3 节论及的凯恩斯就业思想放到一个合适的具体博弈中去考察，并用图形表示出多重均衡。将（国民生产）产出水平（位于纵轴）表示为需求水平（位于横轴）的函数。均衡在国民产出与国民收入相等时取得，即在函数通过 45°线时。何种形状的函数可取得多重均衡？你为什么估计这样的博弈是现实的？假如当前产出水平超过当前收入则收入增长，当前产出水平少于当前收入则收入减少。这样一个动态过程中，哪一个均衡是稳定的，哪一个均衡是不稳定的？

10. 描述一个你亲眼见到或参加过的策略博弈，要求这个博弈中有大量的参与人，而每个人的支付都依赖于其他参与人及其行动。如有可能，请用图形说明你的博弈。鉴于此类博弈大多有无效率结果，请讨论该博弈的现实结果。你在其中发现了这类结果存在的证据吗？

【注释】

[1] 有关公共物品的深入介绍可参见公共经济学方面的教科书，比如 Harvey Rosen, *Public Finance*, 4th ed., Chicago: Richard D. Irwin, 1995 或者 Joseph Stigliz, *Economics of the Public Sector*, 2nd ed., New York: Norton, 1988。

[2] 本章习题 2、习题 3 和习题 5 构造了一些非线性且有多重均衡情况的例子。若想知道这类均衡的一般理论分析，可参考 Thomas Schelling, *Micromotives and Macrobehavior*, New York: Norton, 1978, Chap. 7。若允许参与人的选择为连续性的（例如参与建造的时间长度）而不是只有参与或退出的二元选择，则该理论还可深入。许多关于集体行动的专门著述中讨论了很多这样的情况，例如 Todd Sandler, *Collective Actions: Theory and Applications*, Ann Arbor: University of Michigan Press, 1993; Richard Cornes and Todd Sandler, *The Theory of Externalities, Public Goods, and Club Goods*, 2nd., New York: Cambridge University Press, 1996。

[3] 本段引述的经典巨著已经以不同版本再版过许多次。我们逐一列出每本书的原始出版年份，以及容易获得的再版信息。在每本书中，再版的编辑简单总结介绍了其主要思想。Thomas Hobbes, *Leviathan, or the Matter, Form, and Power of Commonwealth Ecclesiastical and Civil*, 1651, Everyman Edition, London: J. M. Dent, 1973. David Hume, *A Treatise of Human Nature*, 1739, Oxford: Clarendon Press, 1976. Jean-Jacques Rousseau, *A Discourse on Inequality*, 1755, New York: Penguin Books, 1984. Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, 1776, Oxford: Clarendon Press, 1976。

[4] Mancur Olson, *The Logic of Collective Action*, Cambridge: Harvard University Press, 1965.

[5] Garrett Hardin, "The Tragedy of the Commons," *Science*, Vol. 162, 1968, pp. 1243 - 1248.

[6] 这一领域的卓越文献来自 Michael Taylor, *The Possibility of Cooperation*, New York: Cambridge University Press, 1987; Elinor Ostrom, *Governing the Commons*, New York: Cambridge University Press, 1990; Matt Ridley, *The Origins of Virtue*, New York: Viking Penguin, 1996。

[7] 需要达成合作的问题及其解决方案并非人类社会独有。生物学家用基因优势和本能进化的术语解释了动物王国合作行为的一些例子。更多信息参见本书第 13 章和 Ridley 的《德性的起源》(*The Origins of Virtue*)。

[8] 见 Ridley, *The Origins of Virtue* Chaps, 6 and 7。

[9] 社会科学并没有精确和广泛地界定习俗(custom)与规范(norm)等术语,所以此类术语之间的差别并不总是很清楚的。本节对某些术语的界定可能跟其他著作不同。我们的做法类似下文:David Kreps, "Intrinsic Motivation and Extrinsic Incentives," *American Economic Review*, *Papers and Proceddings*, vol. 87, no. 2 (May 1997), pp. 359 - 364。但是 David Kreps 使用规范(norm)一词统称的所有概念在本文中则被加以了分类。

社会学家对规范的分类跟经济学家不同。社会学家认为,像餐桌礼仪这类琐碎规矩称为民俗(folkways),而比较重要的规矩称为民德(mores),正式的规则则称为法律(law)。社会学家也注重价值观(value)与规范之别。而个人价值观跟社会目标的冲突产生于规范而非习俗(conventions)。参见 Donald Light and Suzanne Keller, *Sociology*, 4th ed., New York: Knopf, 1987, pp. 57 - 60。

[10] 进化生物学家 Lee Dugatkin 在其对合作的出现展开研究的著作 *Cheating Monkeys and Citizen Bees* (New York: Free Press, 1999) 中,称这种情况为"自私的协作"。他指出,这类行为在发生危机时更容易出现,因为在那时人人都很关键。在一场危机中,即便只有一个人未能为团体走出危局努力,团体互动的结果也可能是灾难性的。因而每个人在他人努力时自己也会努力。

[11] 关于这种利他惩罚本能,可参见 Ernst Fehr and Simon Gachter, "Altruistic Punishment in Humans," *Nature*, Vol, 415, 2002, pp. 137 - 140。

[12] 参见 Ostrom, *Governing the Commons*, p. 35。本章对规范及制裁的区分类似 Kreps 对规范之功能(ⅲ)和(ⅳ)的区分(Kreps, "Intrisic Motivation and Extrinsic Incentives," p. 359)。

[13] Assar Lindbeck, "Incentives and Social Norms in Household Behavior," *American Economic Review*, Vol 87, no. 2, 1997, pp. 370 - 377.

[14] Orlando Figes, *A People's Tragedy*: *The Russian Revolution1891 - 1924*, New York: Viking Penguin, 1997, pp. 89 - 90, 240 - 241, 729 - 730。也可参见 Ostrom, *Governing the Commons*, p. 23,提供了外部政府试图解决公共资源问题的做法实际上使事情变得更糟糕的另外一些例子。

[15] Avner Greif, "Culture Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies," *Journal of Political Economy*, vol. 102, no. 5, 1994, pp. 912 - 950.

[16] E. E. Schattschneider, *Politics*, *Pressures*, *and the Tariff*, New York: Prentice-Hall, 1935, pp. 285 - 286.

[17] Stephen V. Marks, "A Reassessment of the Empirical Evidence on the Sugar

Program," in *The Economics and Politics of World Sugar Policies*, ed. Stephen V. Marks and Keith E. Maskus, Ann Arbor: University of Michigan Press, 1993, pp. 79 - 108.

[18] David Landes, *The Wealth and Poverty of Nations*, New York: Norton, 1988。第 3 章、第 4 章提供了有关这种效应的一些生动活泼的例子。

[19] 术语"正反馈"可能给人留下这是好事的印象，但是从技术性的语言看，这个术语仅仅刻画了过程而不包括对结果的价值判断。在本例中，同样的正反馈机制既可以导致全部使用 Unix 操作系统也可以导致全部使用 Windows 操作系统；而其中一个结果比另一个结果糟糕。

[20] 并不是每个人都认为 Dvorak 键盘和 Betamax 视频录像机明显地比其他同类产品好。参见如下两篇论文：S. J. Liebowitz and Stephen E. Margolis, "Network Externality: An Uncommon Tragedy," *Journal of Economic Perspectives*, vol. 8, 1994, pp. 146 - 149; "The Fable of Keys," *Journal of Law and Economics*, vol. 33, 1990, pp. 1 - 25。

[21] John Maynard Keynes, *Employment, Interest, and Money*, London: Macmillan, 1936。更正式的失业均衡博弈论模型参见 John Bryant, "A Simple Rational-Expectations Keynes-type Model," *Quarterly Journal of Economics*, Vol. 98, 1983, pp. 525 - 528。

[22] 若要从社会心理学观点全面了解吉诺维斯命案以及对此类情形的分析，可参见 John Sabini, *Social Psychology*, 2nd ed., New York: Norton, 1995, pp. 39 - 44。我们的博弈论模型则基于 Thomas Palfrey and Howard Rosenthal, "Participation and the Provision of Discrete Public Goods," *Journal of Public Economics*, vol. 24, 1984, pp. 171 - 193。

[23] 考虑 $B=10$，$C=8$ 的情况，则 $N=2$ 时有 $P=0.8$，当 $N=100$ 则 P 增加到 0.998，若 N 继续增加则 P 值将逼近 1。当 N 从 2 向正无穷变化，任意某一人采取行动的概率 $1-P$ 就从 0.2 下降到 0。

[24] 同样假设 $B=10$，$C=8$，结果意味着 N 从 2 增加到正无穷会导致无一人实施帮助的概率从 0.64 增加到 0.8，至少有一人帮助的概率从 0.36 下降到 0.2。

[25] John Tierney, "The Boor War: Urban Cranks, Unite-Against Uncivil Behavior, Eggs Are a Last Resort," *New York Times Magazine*, January 5, 1997.

第 13 章　进化博弈

迄今为止，我们研究了具有许多不同特征的博弈——同时行动的与序贯行动的，零和支付的与非零和支付的，操纵博弈规则的策略性行动，一次性的与重复进行的，以及甚至是其中有许多人同时行动的集体行动博弈。然而，在所有的讨论中有一个不变的基本规定——在所有这些博弈中的所有参与人都是理性的：每个参与人都有着内在一致的价值判断尺度，能够计算其在不同策略选择下的后果，并做出最有利于其自身利益的选择。

在应用这种规定时，我们仅仅沿袭了博弈论在大多数情况下所采用的方法。这些方法主要是由经济学家们发展起来的。经济学是建立在理性行为和均衡的双重假定之上的。的确，这些假定在博弈论中也被证明是很有用的。我们对其中参与人通过经验而知道什么是他们的最优选择且因此充分有规律地参与的博弈有了相当的理解。这些假定确保参与人不会错误地在对手面前显得天真无知而被对手所利用。理论也为参与人应该如何玩他们的博弈给出了一些引导。

然而，其他的社会科学家们对理性假设持有极大的怀疑，进而对建立在这一假设基础之上的博弈论也提出了质疑。经济学家们，也正如第 5 章中指出的那样，不应当把理性假设视为理所当然的。麻烦的是需要找到一种可行的替代性假设。尽管我们不能把有意识的和完全可计算的理性强加于参与人，我们也不想放弃这样的一种思想，即某些策略比起其他的策略来说要好一些。我们要用较高的支付去奖赏好的策略；我们要参与人去对成功者进行观察或模仿，并用新的策略去进行尝试；当参与人从玩博弈中获得经验的时候，我们要求好的

策略能够被经常地使用，而不好的策略要更少地被使用。我们在生物进化论和进化动力学中找到了一种对理性假设的可能的替代，于是我们就将本章用于这方面的研究。

13.1 研究框架

生物学中的进化过程为社会科学家所使用的博弈论提供了一种特别有吸引力的参考。进化论的三个基础性假定是：异质性（heterogeneity）、适应性（fitness）和选择（selection）。出发点就是许多动物行为是由基因决定的；一个或几个基因（genotype）的复合体（遗传型）控制着特定的行为模式，称为行为的**表现型**（Phenotype）。基因库的自然多样性保证了种群中表现型的异质性。某些行为比其他的行为更能适应当前的环境，并且表现型的成功可以被量化为**适应性**。人们通常习惯于用众所周知的但是是误导性的词汇"适者生存"去进行思考；然而，对生物学适应性的最终检验不仅要看生物能否成功地生存，而且还要看生物能否成功地繁衍后代。那就是使得动物能够把它的基因传给下一代并且使其表现型能够世代传承的东西。于是，在下一代中更具适应性的表现型在数量上比起较不具有适应性的表现型来说就更加多一些。这种选择过程是动态的，它改变着遗传型和表现型的组合结构，并且也许最终会达到一种稳定状态。

随着时间的推移，偶然的因素会带来新的基因变异（mutations）。许多的这类变异产生了不适应环境的行为（即表现型），它们最终将消亡。但是偶尔会有一种变异会带来一种较为适应环境的新的表现型，于是这样的一种变异基因就会成功地侵入（invade）一个种群，也就是说，它将扩散开来并且成为种群的一个重要的组成部分。

在任何时候，种群可包含其一些甚至其全部的生物学上可想象到的表现型。那些比其他表现型更具适应性的表现型在数量上将会增加，一些不具适应性的表现型会消失，而且其他目前在种群不存在的表现型会尝试着入侵它。当种群不能被任何变异成功地侵入时，生物学家就称该种群结构及其当前表现型是**进化稳定**（evolutionary stable）的。这是一种静态检验，但经常使用的是一种更加动态的标准：从种群中任何一个表现型的复合体出发，如果某种群结构是动态选择的极限结果，那么该结构就是进化稳定的。[1]

表现型的适应性取决于个体或有机体与环境的关系，例如，一种特定鸟类的适应性就取决于它们翅膀的空气动力学特征。它同时还取决于存在于环境中的不同表现型的比例的复杂性——相对于整个物种中其余的种群来说，它的翅膀利用空气动力的程度。因此，一种特定动物的适应性以及它的行为特点，诸如攻击性和社会性，取决于种群中其他大多数成员是攻击性的还是温顺的，是群居的还是独行的，等等。从我们的研究角度来说，一种物种内部的不同表现型之间的这种相互作用是最令人感兴趣的方面。当然，有时一种物种中的个体也与其他物种中

的成员发生相互作用，于是，比如说，一种特定类型的羊的适应性就取决于当地狼群的现有特征。我们也考虑这种类型的相互作用，但是仅仅是在我们研究完种群内部的情形之后。

生物进化过程发现了与博弈论中十分相似的概念。一种表现型的行为可被看做动物在与其他动物相互作用中的一种策略——比如，是攻击还是退却。区别是这种策略的选择并非像在标准博弈论中那样是一种有目的的计算，相反，它是表现型的遗传性的先天本能。相互作用带给表现型支付。在生物学里，支付衡量进化或者繁殖的适应性。当我们在生物学之外应用这种思路时，它们在一些颇具争议的社会博弈、政治博弈和经济博弈中的成功便被赋予了新的意义。

支付或适应性的数值可表示在恰如标准博弈中那样的支付表中，某种动物的可想象出来的所有表现型按矩阵的各行排列，而其他动物的则沿矩阵的各列排列。倘若有更多的动物同时相互作用——在生物学中被称为**领地博弈**（playing the field），支付就可用第 12 章中的集体行动博弈中的那种函数表示。我们将在本章的大多数地方考虑成对之间的匹配，并在 13.9 节简要地考察其他情形。

因为种群是表现型的一种混合体，从中挑选出来的不同对表现型会给它们的互动带来不同的策略组合。表现型适应性的实际定量测度是其在与种群中其他表现型之间的所有相互作用中获得的平均支付。那些具有较强适应性的动物会有更大的进化成功机会。种群动态变化的最终结果将会是一种进化稳定的种群结构。

生物学家们非常成功地运用了这一方法。攻击性行为与合作行为的组合，巢穴的选些及其他无法用传统观点解释的更多的现象都可以理解为选择更具适应性的策略的一种进化过程的稳定结果。有趣的是，生物学家在使用已有的博弈论框架的时候发展了进化博弈的思想，他们沿用了博弈论的语言，但修改了参与人有意识地最大化以满足自身需求的假设。目前，博弈论专家们又反过来运用从生物进化博弈研究中所得到的启示来丰富他们自己的学科。[2]

的确，进化博弈理论放松了理性行为的假定，为新的博弈理论研究方法提供了一个现成的框架。[3]根据进化博弈的这种观点，参与人并没有自己选择策略的自由，有的参与人天生就得选择某种策略，而其他参与人天生就会选择另外的策略。策略的遗传性思想在博弈论中可比在生物学中获得更加广泛的解释。在人类的相互作用中，有许多理由使得策略是被嵌入到参与人的大脑里的——不仅仅是由于遗传，还由于（并且可能是更加重要的）社会性、文化背景、教育及基于以往经验的归纳性法则。一种人群可能由具有不同背景或经验的不同个体的复合体组成，而这些个体先天性地被嵌入了不同的策略。所以有的政治家宁可放弃选举成功也要坚决地坚守一定的道德或伦理规范，而另外的政治家却更关心他们自己能否连任；类似地，某些公司只追求利润，而另外的公司追求的却是社会和生态的目标。我们可将那些按照这样的方式被牢牢固定下来的每一个可想象出来的策略称为这里所研究的参与人群体的表现型。

从具有先天性策略的异质性的种群中随机挑选出若干对表现型的组合与同一物种或不同物种的其他表现型组合重复地相互作用（进行博弈）。每次相互作用中，参与人的支付取决于双方的策略；这种依赖性由通常的"博弈规则"所支配

并且在博弈表或博弈树中得到反映。我们可以将某一策略的适应性定义为该策略与种群中所有其他策略博弈时的平均支付或总支付。某些策略相比于其他策略具有更高水平的适应性；在种群的下一代即下一轮的博弈中，那些适应性更强的策略将会被更多的参与人使用并繁殖扩散开来；适应性较低的策略则只有较少的参与人才使用，然后就失去活力且逐渐消失。有时，某些个体会试试或者采用过去没有使用过的策略，而这些策略是逻辑上可想象出来的。这就对应于变异的产生。倘若这样的新策略要比当前采用的策略更具适应性，它就会开始被群体中更大比例的个体所采用。核心的问题是种群中特定策略的选择性繁殖扩散、失去活力且消失以及变异的过程是否有一个进化稳定的结果，并且如果是这样，稳定的结果又是什么。对于我们刚才谈及的例子，社会是否最终会变成所有政治家只关心各自能否连任和所有企业都只关心利润的情形呢？本章中我们将提出用于回答这些问题的框架和方法。

尽管我们使用了生物学的类比，但在社会经济博弈中具有较强适应性的策略会繁殖扩散开来以及具有较低适应性的策略会消失的理由与生物学中严格的遗传机制不尽相同：在前一回合中混得好的参与人会将信息传递给下一回合中的朋友或同事们，而在前一回合中混得不怎么样的参与人会对成功的策略进行观察，然后接着去模仿它们，于是在随后的博弈中，参与人开始进行一些有目的的思考和对先前的经验法则的修正。在大多数策略博弈中，这种“社会的”和“教育的”传递机制远比任何生物学遗传显得更为重要；的确，这就是立法者连任目标和企业利润最大化动机如何得以增强的过程。最后，新的策略的有意识的试验代替了生物学博弈中的偶然变异。

生物博弈的进化稳定结构可有两种类型。首先，某个表现型可比任何其他表现型更具适应性，并且种群会趋向于由其单独构成。这样一种进化稳定结果被称为**单态型**（monomorphism），即单一的形式。此时，这种唯一出现的策略被称为是一种**进化稳定策略**（evolutionary stable strategy，简写为 ESS）。另一种可能性是两种或更多种表现型可能具有同样的适应性（并且比其他表现型更具适应性），故它们可以以一定比例共存。这时我们称种群表现出**多态型**（polymorphism），即形式的多样化。在这个种群中，如果没有新的表现型或可行的变异比现有多态型种群中的表现型更具适应性，那么这种状态就是稳定的。

多态型与博弈论中的混合策略概念是密切相关的。但是，也存在一个重要的区别。若要获得多态型，不需要个别参与人采用混合策略。每个参与人都选择纯策略，但是种群却会因不同参与人选择不同纯策略而呈现为一种混合策略。

种群及其可想象的表现型集合，表现型之间相互作用中的支付矩阵，以及在种群中占一定比例的各种表现型对进化的作用（当然，这与表现型的适应性相关）构成了一个进化博弈。种群的进化稳定结构即进行博弈的均衡。

在本章中，像通常那样，我们将通过一系列的说明性例子来提出某些思路。我们从对称博弈开始，其中两个参与人的地位是相同的——比如，同一物种中的两个成员为争夺食物或配偶而相互竞争；在社会科学的理解里，他们可以是为了在公共事务方面继续行使权力而竞争着的两个候选官员。根据博弈的支付表，每

个参与人都可以被安排成行参与人和列参与人而并不影响最后的结果。

13.2 囚徒困境

假设一个种群由两种表现型组成。一种由那些天生就是合作者的参与人组成，他们的行为总是使得所有参与人联合最优的结果有可能出现。另一种由背叛者组成，他们的行为总是只顾及他们自己。作为一个例子，我们利用第 5 章描述过的餐馆定价博弈。关于这个例子，我们在第 11 章中曾经给出了一个简化的版本。在这里，我们使用这个简化的版本，其中只有两种定价策略可供选择，即联合最优价格 26 美元或纳什均衡价格 20 美元。合作性的餐馆老板会总是选择 26 美元，而背叛性的餐馆老板将总是选择 20 美元。表 13—1 给出了这种离散型囚徒困境的一次性博弈中每一种表现型的支付（利润），它是由表 11—2 得到的。在这里，我们简单地将参与人称为“行”和“列”，每个参与人都是从种群中随机挑选出来的。

表 13—1　　定价的囚徒困境（百美元/月）

		列	
		20(背叛)	26(合作)
行	20(背叛)	288，288	360，216
	26(合作)	216，360	324，324

回忆一下，在进化的情形里，没有人会在背叛与合作之间进行选择，每一个人都是“先天地”被赋予这一种或另一种天生的个性特点。哪一种会是更为成功的（更具适应性的）个性特点呢？

一个背叛型的餐馆老板如果与另一个背叛型的餐馆老板相遇，则得到 288 的支付（每月 28 800 美元）；当与一个合作型的相遇时，则得到 360(每月36 000美元）的支付。合作型的与背叛型的相遇时，获得支付 216(每月 21 600 美元)；而当与另外一个合作型的相遇时，获得支付 324(每月 32 400 美元)。无论与哪种类型的对手相遇，背叛型的都比合作型的要好。[4] 因此无论种群中两种类型的比例如何，背叛型的比起合作型的来说都有着更高的期望支付（因而也更具适应性）。

稍微正规一点，设合作者的比例为 x。考虑任何一个特定的合作者。在一次随机抽取中，他遇到另外一个合作者（且得到 324）的概率为 x，而遇到一个背叛者（且得到 216）的概率为 $(1-x)$。因此，一个典型合作者的期望支付为 $324x+216(1-x)$。对于一个背叛者，遇到一个合作者（且得到 360）的概率为 x，而遇到另外一个背叛者（且得到 288）的概率为 $(1-x)$。所以一个典型背叛者的期望支付为 $360x+288(1-x)$。现在，立即就会有

$$360x+288(1-x)>324x+216(1-x)$$

对所有在 0 和 1 之间的 x 都成立。

所以，背叛者有比较高的期望支付并且比合作者更具适应性。这将会导致背叛者的比例一代一代地增加（x 减少），直到整个种群都全由背叛者组成。

如果种群最初全部由背叛者组成，会有什么事情发生呢？在这样的情形里，则不会有变异（试验性的）性合作者生存下来并增加繁衍到取代整个种群；换句话说，变异性合作者不可能成功地侵入由背叛者组成的种群。即使 x 很小，即种群中合作者的比例很小，合作者的适应性仍不如现有的背叛者，并且其种群比例将不会上升而只会下降到 0；因而这种合作型变异品种就会消亡掉。

我们的分析表明背叛者比合作者有更强的适应性以及由背叛者组成的种群不能被变异性合作者侵入。因此，种群的进化稳定结构是单态型的，由单个的策略或表现型组成。我们因而称背叛为这种囚徒困境博弈中种群的进化稳定策略。要注意的是，背叛是同一博弈中的理性行为分析里的严格占优策略。该结果是非常具有一般性的：如果一个博弈有严格占优策略，那么该策略也是 ESS。

13.2.1 重复博弈的囚徒困境

我们在第 11 章中已经看到重复进行的囚徒困境博弈是如何使得理性的参与人为了双方的利益而有意识地维系合作的。让我们来看看在进化博弈中是否也存在类似的可能。假设被选出来的每一对参与人接连不断地进行三轮的囚徒困境博弈。每个参与人从这种互动中所得的总支付为他在三个回合中得到的支付之和。

每个个体仍然是既定地只使用一个策略，但这个策略必须是一种完整的行动计划。在有三个回合的博弈中，一个策略对第二个回合或第三个回合中的行动规定取决于第一个回合或第二个回合里所发生的情况。比如，“无论如何我都选择合作”，“无论如何我都选择背叛”都分别可以是一个策略。但是“我开始选择合作；如果你在前面回合合作的话，我将继续合作；如果你在前面的回合背叛的话，我就会在随后一直选择背叛”也是一个策略。事实上，最后的策略正好就是以牙还牙策略（tit-for-tat，简写为 TFT）。①

为了使得初步的分析简单化，我们在本节假设种群中只可能存在两种类型的策略：总是背叛（A）和以牙还牙（T）。从种群中随机挑选出来几对个体，每一对被挑选出来的个体都进行特定次数的博弈。每一个参与人的适应性就是其在与特定对手进行的所有重复性博弈中得到的支付之总和。我们来考察当每一对个体进行两次、三次以及更为一般性的 n 次这样的重复博弈时，将会发生些什么。

① 疑原书有误，应为冷酷策略。 ——译者注

1. 两次重复

表 13—2 给出了一个博弈的支付表，其中餐馆老板群体中的两个成员相遇且博弈两次。倘若两个参与人都是 A 类型的，则两个都背叛两次，并且表 13—1 表明每个人每一次都得到 288，总数为 576。如果两个参与人都是 T 类型的，则背叛就不会出现，每个人每一次都得到 324，总数为 648。如果一个是 A 类型的而另外的一个是 T 类型的，则在第一回合中，A 类型的背叛而 T 类型的合作；所以前者得到的是 360 而 后者得到的是 216。在第二个回合中，双方都背叛并且得到 288。故 A 类型的总支付为 360＋288＝648，且 T 类型的总支付为 216＋288＝504。

表 13—2　　两次重复的囚徒困境中的结果（百美元/月）

		列	
		A	T
行	A	576，576	648，504
	T	504，648	648，648

在两次重复的囚徒困境中，我们看到 A 类型仅仅是弱占优的。不难看出，如果种群中全部都是 A 类型，则 T 类型变异不能侵入，且 A 类型就是一个 ESS。但是，如果种群中全部都是 T 类型，则 A 类型变异不会比 T 类型做得更好。这是不是就意味着 T 类型就一定是另外的一个 ESS，正如在这个博弈的理性博弈框架分析中它会是一个纳什均衡那样呢？答案是否定的。倘若种群最初都是由 T 类型组成的，并且有少量的 A 类型变异进入，则变异会在大多数时间里与占优势的 T 类型相遇，并进行对局。但是，有时 A 类型变异会与另外的 A 类型变异相遇，并进行对局，当然此时的 A 类型变异比 T 类型（当 T 类型与 A 类型对局时）更占优势。因此，变异具有比占优势的表现型稍微强一些的适应性。这一优势就给种群中的变异在比例上带来一种增加，虽然只是慢慢地增加。所以，全部由 T 类型组成的种群会被 A 类型变异逐渐地侵入，T 类型不是一个 ESS。

我们的推理依赖于关于 ESS 的两种检验。首先我们看看当与占优势的表现型相遇时，变异比占优势的表现型做得是更好还是更差。倘若这个基本的准则给出一个清晰的答案，问题就解决了。如果这个基本的准则给出的是一个不明确的回答，则我们就要用另一种方法去明确它，或者说是第二个准则：如果与变异相遇，变异比占优势的表现型做得是更好呢还是更差？存在不明确的答案的情形是罕见的，大多数时候我们并不需要第二个准则，但正是在那里存在着如表 13—2 所说明的那种情形。[5]

2. 三次重复

现在假定来自（A，T）种群里的每一个匹配对将进行三次博弈。表 13—3 给出了每一种类型参与人在与每一种类型的对手相遇时的适应性结果，它是三次相遇的支付的加总。

表 13—3　三次重复的囚徒困境中的结果（百美元/月）

		列	
		A	T
行	A	864，864	936，792
	T	792，936	972，972

为了理解这些适应性数值是如何出现的，考虑两种情况。当两个 T 类型参与人相遇时，两人开始是合作的，而且因此在第二次两人也合作以及在第三次也是如此；两人每次都得到 324，三个月每人总共得到 972。当一个 T 类型参与人与一个 A 类型参与人相遇时，后者开始做得很好（A 类型参与人得到 360，T 类型参与人得到 216），但是随后 T 类型参与人也在第二次和第三次背叛，并且每人都在这两次中得到 288（A 类型参与人的总支付为 936，T 类型参与人的总支付为 792）。

两种类型的相对适应性依赖于种群的构成。如果种群整个都几乎是 A 类型的，则 A 类型要比 T 类型更具适应性（因为 A 类型在大多数时间会与 A 类型相遇且得到 864，而 T 类型则经常地获得 792）。但是，倘若种群整个都几乎是 T 类型的，则 T 类型要比 A 类型更具适应性（因为 T 类型经常地遇到 T 类型得到 972，但 A 类型在这样的场合得到的是 936）。每一种类型在其已经在种群里占优势时是更具适应性的。所以，T 类型在种群都是 A 类型时是不能成功侵入的，反之亦然。现在，种群存在着两种可能的进化稳定结构：在一种结构中，A 类型是 ESS；而在另一种结构里，T 类型是 ESS。

接着再考虑当初始种群由两种类型混合而成时的进化动态过程。种群的构成将如何随时间而进化？假定种群中有 x 比例的是 T 类型，余下的 $(1-x)$ 是 A 类型。[6]一个 A 类型参与人，在与从这样一个种群里挑选出来的各种各样对手的竞争中，当与 T 类型参与人遭遇时获得 936，这种情形在 x 比例的时间里发生，而在与 A 类型参与人遭遇时获得 864，这种情形在 $(1-x)$ 比例的时间里发生。这样就给出每一个 A 类型参与人的平均期望支付：

$$936x+864(1-x)=864+72x$$

类似地，一个 T 类型参与人的平均期望支付为

$$972x+792(1-x)=792+180x$$

那么，一个 T 类型参与人会在获得更多的平均期望支付情况下比一个 A 类型参与人更具适应性，即

$$792+180x>864+72x$$
$$108x>72$$
$$x>2/3$$

换句话说，如果种群中有 2/3 以上都是 T 类型参与人，则 T 类型参与人是更具适应性的，且其所占的比例会增加，直到达到 100%为止。如果种群开始时 T 类型参与人少于 2/3，则 A 类型参与人会更具适应性，且 T 类型参与人的比例会下降直

到为 0 或者 A 类型参与人的比例为 100%为止。进化动态过程使得种群移向两种极端情形之一，两者都是可能的 ESS。动态过程带来了与变异入侵的静态检验相同的结论。这是进化博弈的一个一般的——虽然并不是放之四海而皆准的——特征。

这样，我们已经识别出了种群的两种进化稳定结构。在每一种情形中，种群都由一种类型组成（单态型的）。例如，倘若种群最初是由 100%的 T 类型组成，则即使在少量的 A 类型出现之后，种群复合体会仍然至少有 66.66%的 T 类型；T 类型仍然是更具适应性的类型，而且变异 A 类型将消亡。类似地，如果种群最初是由 100%的 A 组成，则少量的 T 类型变异会使得种群复合体中只有不超过 66.66%的 T 类型；所以 A 类型会更具适应性，且变异 T 类型将消亡。而且，恰如我们将在后面小节中看到的，类型 N 的实验性变异绝不能够在或者主要是 T 类型或者主要是 A 类型的种群中取得成功。

如果种群最初恰好有 66.66%的 T 类型参与人（并且有 33.33%的 A 类型参与人），则会发生什么呢？这时两种类型具有同样的适应性。我们可将其称为多态型。但它实际上并不是一种进化稳定结构的恰当的候选者。种群仅仅能够在任何一种类型的变异出现之前维持这种微妙的平衡。有时，这种变异迟早必定会发生。这种变异的到来将使得适应性的计算反过来变得有利于变异的类型，且其优势会不断积累直到达到 100%的那种类型的 ESS 为止。这正是进化稳定的第二准则的一个应用。我们有时不严谨地把这样一种结构说成是一种非稳定均衡，这样就可以与通常的博弈论保持一种对应。在通常的博弈论中，变异并没有被考虑进去，并且微妙的平衡均衡是可以持续下去的。但是依照生物学过程的严格逻辑，它根本就不是一个均衡。

这种推理可用一种简单的图形来说明，它十分像我们在计算理性参与人混合策略均衡中的均衡比例时所画的那幅图。唯一的区别就是，在进化的场合里，其中分离性策略的比例并不是一个任何参与人的选择问题而是整个种群的一种性质，如图 13—1 所示。沿着水平轴，我们用 0 到 1 之间的数度量种群中 T 类型参

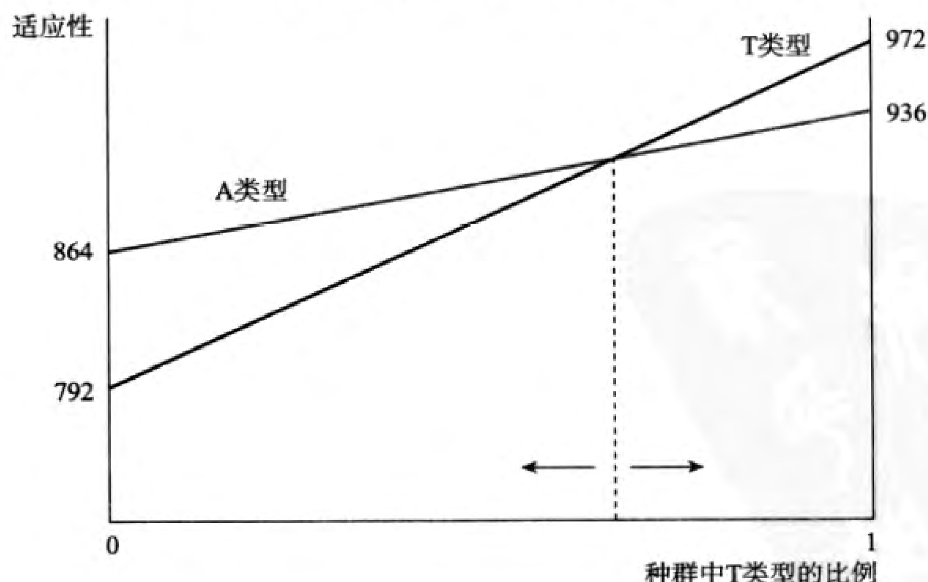


图 13—1　三次重复囚徒困境博弈的适应性及均衡

与人的比例 x。我们沿纵轴度量适应性。每一条线都表示一种类型的适应性。T 类型的线开始时比较低（位于 792，与 A 类型线的 864 形成对比）并且终于较高的位置（972，与 A 类型线的 936 相对）。两条线在 $x=2/3$ 处相交。在该点的右端，T 类型更具适应性，所以其种群比例随着时间而增加并且 x 增加到 1。类似地，在该点的左端，A 类型更具适应性，故其种群比例随着时间而增加并且 x 下降到 0。这样的图形在直观上是很有用的，我们将广泛地使用它。[7]

13.2.2 多次重复博弈

倘若每一对个体进行某一非特定次数的重复博弈，会有什么事情发生呢？让我们来集中考虑一个种群，它仅仅由 A 类型和 T 类型所组成，其中随机配对地相互作用进行 n 次（其中 $n \geqslant 2$）。表 13—4 给出了进行 n 次重复博弈的总结果表。当两个 A 类型相遇时，他们总是背叛且每次都得到 288；故每人在 n 次博弈中共获得 $288n$。当两个 T 类型相遇时，他们以合作开始，没有人会首先背叛；所以他们每次获得 324，总共是 $324n$。当 A 类型与 T 类型相遇时，在第一回合里 T 类型合作但 A 类型背叛，所以 A 类型得到 360 且 T 类型获得 216；其后，T 类型在余下的回合里对 A 类型先前的背叛行为进行报复，并且在所有余下的 $(n-1)$ 个回合里每人都获得 288。因此，A 类型在 n 次应对 T 类型过程中总共获得 $360+288(n-1)=288n+72$，而 T 类型在 n 次应对 A 类型过程中总共获得 $216+288(n-1)=288n-72$。

表 13—4　　n 阶段重复的囚徒困境博弈的支付

		列	
		A	T
行	A	$288n$，$288n$	$288n+72$，$288n-72$
	T	$288n-72$，$288n+72$	$324n$，$324n$

倘若种群中 T 类型的比例为 x，则一个典型的 A 类型平均得到 $x(288n+72)+(1-x)288n$，并且一个典型的 T 类型平均得到 $x(324n)+(1-x)(288n-72)$。因此当下面的条件满足时，T 类型就更具适应性：

$$x(324n)+(1-x)(288n-72)>x(288n+72)+(1-x)288n$$

$$36xn>72$$

$$x>\frac{72}{36n}=\frac{2}{n}$$

再一次，我们获得两个单态型的 ESS：一个 ESS 中全部都是 T（或 $x=1$，即从 $x>2/n$ 中任一点起，收敛于 $x=1$）；另一个 ESS 中全部都是 A（或 $x=0$，即 $x<2/n$ 中任一点起，收敛于 $x=0$）。如图 13—1 所示，在平衡点 $x=2/n$ 处还存在一个非稳定的多态型均衡。

要注意的是，平衡点处T类型的比例依赖于n；当n较大时它就比较小。当$n=10$时，它等于2/10，或0.2。所以，如果种群最初有20%的T类型的话，那么博弈重复进行10次后，T类型的比例将增加到100%。回忆一下，当博弈只进行三次时（$n=3$），T类型只需要在初始时占到67%以上的比例就可以达到同样的结果。而在两次重复博弈中，T类型在初始时的比例必须为100%才可避免被淘汰。在x的临界值的表达式中，我们可以看到产生这种结果的原因。它表明，当$n=2$，T类型尚未更具适应性时，x必须等于1。请记住，全部由T类型组成的种群会达成合作。因此，当博弈被重复多次时，较大范围的初始条件将会导致合作行为。从这个意义上讲，重复博弈次数越多，合作的可能性就越大。我们正看到这样的事实：相互作用的时间越长，达成合作的价值越大。

13.2.3 进化博弈模型与理性参与人模型之间的比较

最后，让我们回到曾在表13—3中解释过的三次重复博弈，并且不使用进化模型，而考虑它是在两个有意识地进行理性决策的参与人之间进行的。纳什均衡是什么呢？这里有两个纯策略纳什均衡，其中一个是两者都选A，另外一个是两者都选T。还有一个混合策略均衡，其中在67%时间里选T，在33%的时间里选A。前面的两个恰好就是我们已发现的单态型的ESS，而第三个就是不稳定的多态型进化均衡。换句话说，进化与有意识地进行理性决策的观点之间存在着密切的关系。

这绝非偶然。一个ESS必定是由具有相同支付结构的有意识地进行理性决策的参与人所进行的博弈的一个纳什均衡。为了理解这一点，暂时假定是相反的。如果所有参与人都选择某些策略，比如说就是某一策略S，它不是一个纳什均衡，那么，某个其他的策略，不妨说就是策略R，被参与人用来对付S时必然会带来一个更高的支付。在一个选择S的种群中，A类型变异选择R策略将获得更强的适应性，从而会成功地侵入这个种群。因此，S策略不可能是ESS。换句话说，假如所有参与人都选择S不是一个纳什均衡，则S就不可能是一个ESS。这就相当于说"如果S是一个ESS，那么所有参与人都采用S必定是一个纳什均衡"。

因此，进化方法以一种迂回的方式为理性人方法提供了一种理由。即使当参与人并不是有意识地追求支付最大化的时候，假如更为成功的策略被更频繁地使用，而不太成功的策略逐渐消失，而且整个过程最终收敛于一个稳定的策略，那么，结果就一定与有意识地选择最大化策略时得到的结果相同。

虽然一个ESS一定是对应的理性决策博弈的纳什均衡，但是反过来却并不成立。我们已经看到了两个这样的例子。在表13—2中的理性决策两次重复囚徒困境博弈里，T会是这样一种比较弱的意义上的一个纳什均衡，即如果两个参与人都选择T，没有人会在转而选择A中获得正的收益。但是在进化方法中，A类型能够作为一种变异出现并且能够成功地侵入由T类型所组成的种

群。而且在表13—3和图13—1中的三次重复囚徒困境博弈里，理性决策将生成一个混合策略均衡。但这个混合策略均衡的生物学对应者——多态型的状态——可被变异成功地侵入，因此并不是一个真正的进化稳定结构。所以，生物学的稳定概念有助于我们从理性决策博弈的多重纳什均衡中进行挑选。

我们对重复博弈的分析存在着一种局限。在一开始，我们只考虑了两个策略——A和T，并假定不存在其他的策略或者没有变异出现。在生物学上，出现的变异类型是由所考虑的基因决定的。在社会、经济或政治的博弈中，新的策略的产生大概是由历史、文化以及参与人的经验所支配的，人们吸收、处理信息的能力以及尝试各种各样策略的能力也起着一定作用。然而，我们对特定博弈中可能存在的策略所施加的限制对于决定哪个策略（如果有的话）可能是进化稳定的有着重要意义。在三次重复囚徒困境博弈例子里，如果我们允许这样一种策略S——在第一轮中合作而在第二轮和第三轮里都是背叛，那么S类型的变异就能够成功地侵入全部由T类型组成的种群，所以T策略就不会是一个ESS。我们将在本章末的习题6中进一步讨论这种可能性。

13.3 懦夫博弈

回忆前面提到的20世纪50年代的年轻人驾驶着他们的汽车相向行驶，看看谁会为了避免相撞而首先转向的博弈。现在我们假设参与人在这种事情上是没有选择的：每个人生来就或者是懦弱的（总是转向逃掉）或者是具有大男子气概的（总是径直向前开）。种群由这两种类型混合而成。每周都随机性地挑选出一对参与人玩这种游戏。表13—5给出的就是对于任意两个这样的参与人比如A和B的支付矩阵（数字是我们在表4—14和表8—3中曾经使用过的数字）。

表13—5　　懦夫博弈支付表

		B	
		懦夫	勇士
A	懦夫	0，0	−1，1
	勇士	1，−1	−2，−2

这两种类型将会如何发展变化呢？答案依赖于最初种群中各类型所占的比例。如果种群中几乎全是懦夫，那么勇士的变异将会取胜并且在大部分时间里都得到1分，而所有懦夫在遇到与他们同样类型的人时得到0分。但是，如果种群几乎全部由勇者组成，则懦夫的变异得到−1分，它看起来是很糟糕的，但是却比所有人都是勇士的时候得到−2分要好。你可以用生物学和20世纪50年代的性别歧视方式来恰当地思考这一问题：在一个全是懦夫的种群中，一个勇士的出现会使所有其余的人都显得懦弱，并且会给所有的女孩子都留下深刻的印象。但

是如果种群中全部都是勇士，他们在大多数的时间里都会躺在医院里，女孩们将不得不去找那些少数的健康的懦夫了。

换句话说，种群中那些相对稀少的类型将会更具适应性。因此它能成功地侵入由其他类型组成的种群。我们想看看种群在均衡中的两种类型，也就是说，我们预期有一个混合体的或多态型的 ESS。

为了找出在这个 ESS 中的懦夫与勇士的比例，让我们先来计算一下在一般的混合种群中不同类型的适应性。记 x 为勇士在种群中所占的比例，从而 $(1-x)$ 为其中懦夫所占的比例。一个懦夫在 $(1-x)$ 比例的时间内与另一个懦夫相遇并得到 0 支付，在 x 比例的时间里遇到一个勇士并得到支付 -1。因此懦夫的适应性为 $0\times(1-x)-1\times x=-x$。类似地，勇士的适应性为 $1\times(1-x)-2x=1-3x$。如果下面的条件成立，则勇士就更具适应性：

$$1-3x>-x$$
$$2x<1$$
$$x<\frac{1}{2}$$

如果种群中勇士的数量少于一半，那么勇士就更具适应性并且他们的比例会上升。另一方面，如果种群中勇士的比例大于一半，那么懦夫将会更具适应性并且勇士的比例会下降。不管怎样，种群中勇士的比例都会趋向 1/2，并且这个 50%对 50%的混合体将是稳定的多态型的 ESS。

图 13—2 中给出了这一结果的图像。每一条直线都显示了一种类型的适应性（在与种群中的成员随机性配对的情况下的期望支付），此时勇士的比例为 x。懦夫的适应性与勇士比例的函数关系是 $-x$，正如我们在前面所看到的那样，这是一条缓慢下降的直线，它从 $x=0$ 时高度为 0 的地方出发，并且在 $x=1$ 时趋于 -1。勇士对应的函数为 $1-3x$。这是一条迅速下降的直线，它从 $x=0$ 时高度为 1 的地方出发，然后在 $x=1$ 时下降到 -2。当 $x<1/2$ 时，勇士的直线位于懦夫的直线的上方；而当 $x>1/2$ 时，勇士的直线就位于懦夫的直线的下面。这表明当 x 很小时，勇士就更具适应性，当 x 较大时，懦夫的适应性就更强。

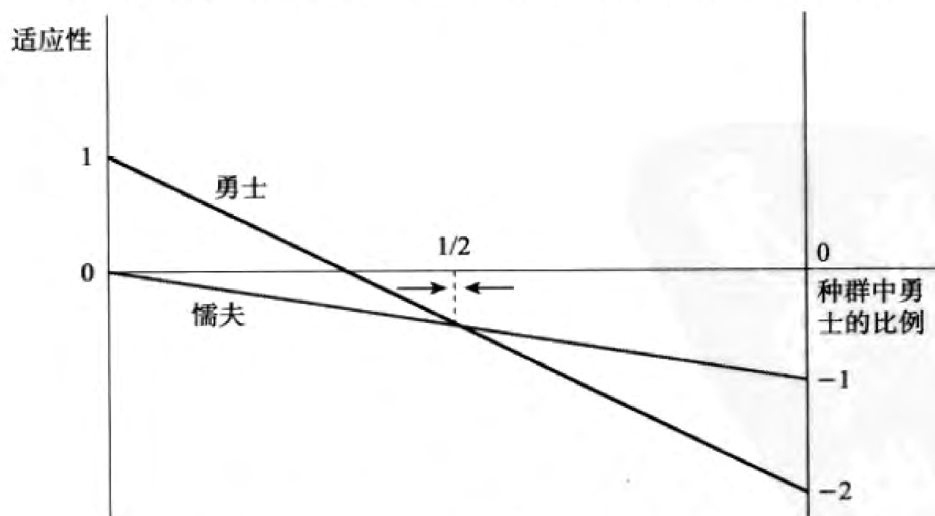


图 13—2　懦夫博弈的适应性和多态型均衡

现在我们可以把这个博弈的进化理论与早先在第 4 章和第 8 章中基于参与人都是有意识进行理性策略计算这一假定上的理论进行比较。在那里我们发现有三个纳什均衡：两个是纯策略的，其中一个参与人径直向前而另外的一个转向；还有一个混合策略的，其中每一个参与人都以 1/2 的概率径直向前，以 1/2 的概率转向。

如果种群全部由勇士组成，则所有的参与人都具有同样的适应性。类似地，一个全部都是懦夫的种群中，所有参与人也都具有同样的适应性。但这些单态型结构是不稳定的。在一个全部由勇士组成的种群中，一个懦夫的变异将会比勇士更具适应性从而成功地侵入并存活下来。[8]

一旦一些懦夫站稳了脚跟，无论他们数量多么少，我们的分析表明，他们的比例都会不可阻挡地上升到 1/2。类似地，一个全部由懦夫组成的种群也易于受到勇士变异的成功侵入，并且这个过程再次向着同样的多态型进化。因此，多态型结构是唯一真正的进化稳定结果。

最为有趣的是理性博弈中的混合策略均衡与进化博弈中的多态型均衡之间的联系。前者均衡策略中的混合比例与后者的种群中不同表现型（类型）所占的比例完全相同，即 50%的懦夫和 50%的勇者。但是在理解上却有所不同：在理性框架中，每个参与人混合运用他自己的策略；而在进化框架中，每个个体使用的是一种纯策略，但是不同类型的个体使用的是不同的策略，所以我们在种群中所见到的是一种混合体。[9]

理性博弈中的纳什均衡与具有相同支付结构并按照进化规则进行的博弈的稳定结果之间的对应关系是一个非常一般的命题，并且随后在 13.6 节中我们还将进一步看到它的一般性。事实上，进化稳定性为我们在这样的理性博弈中的多重纳什均衡中选择一个均衡提供了一个额外的理由。

当我们从理性角度看待懦夫博弈的时候，混合策略均衡看起来似乎令人费解。它可能导致很大的错误。每一个参与人在一半的时间里驾驶着车径直向前，所以在四分之一的时间里他们就会发生一次相撞；而纯策略均衡就避免了这种相撞。在那个时候你会感到混合策略是多么地令人不快了，甚至会对我们花如此多的时间在这上面感到奇怪。现在你就看到其中的道理了。这个看起来很奇怪的均衡是作为一种自然动态过程的稳定结果而出现的，在这一过程中，每个参与人在应对所遇到的种群时，都试图改善他的支付。

13.4 安全博弈

关于第 4 章介绍的策略博弈重要类别，我们已从进化博弈的角度研究了囚徒困境和懦夫博弈，现在还剩下安全博弈。我们曾在第 4 章中用两个大学生哈里与萨莉决定去哪里喝咖啡的故事来说明这类博弈。在进化理论里，每个参与人生来就喜欢星巴克或者是本地咖啡店，并且种群里包括了两种类型的参与人。我们在这里假定，这两种类型的配对——我们一般把这两种类型称为男人和女人，是每

天随机地挑选出来进行博弈的。我们现在用 S(星巴克) 和 L(本地咖啡店) 表示策略。表 13—6 表示该博弈中随机配对的支付表，支付与前面在表 4—12 中说明的一样。

表 13—6　　　　安全博弈支付矩阵

		女人	
		S	L
男人	S	1，1	0，0
	L	0，0	2，2

如果这是一个由理性地选择策略的参与人所进行的博弈，则会有两个纯策略均衡：(S，S) 与 (L，L)。对于两个参与人来说，后者都要好一些。如果他们直接进行通信和协调，可以十分容易地搞定。但是，倘若他们是独立地进行决策的话，他们就需要通过预期收敛——也就是说，通过找到一个聚点——来协调。

理性决策博弈有第三个均衡，是混合策略的，我们在第 8 章里就发现了。在那个均衡中，每个参与人以概率 2/3 选择星巴克，以概率 1/3 选择本地咖啡店；每个参与人的期望支付都是 2/3。恰如我们在第 8 章中所给出来的那样，这时的期望支付比他们达到稍欠吸引力的纯策略均衡 (S，S) 时所得到的支付少，因为独立地使用混合策略使得参与人在许多时候都做出冲突的或者是糟糕的选择。这里，糟糕的结果（支付为 0）发生的概率是 4/9——两个参与人都几乎在一半的时间里去了不同的约会地点。

如果是进化博弈的话，会出现什么情况呢？如果种群很大，每一个成员都有先天性的行为，或者是选择 S，或者是选择 L。从种群中随机选择出来两个人，要求他们约会一次。

假定 x 是种群中 S 类型的比例，则 $(1-x)$ 就是 L 类型的比例。那么，S 类型的适应性——随机遇到某一类型的期望支付——是 $x\times1+(1-x)\times0=x$。类似地，每一个 L 类型的适应性是 $x\times0+(1-x)\times2=2(1-x)$。因此，在 $x>2(1-x)$ 或 $x>2/3$ 时，S 类型就具有更强的适应性。在 $x<2/3$ 时，L 类型具有更强的适应性。在平衡点 $x=2/3$ 处，两种类型具有同样的适应性。

正如在懦夫博弈中那样，理性决策下的混合策略均衡中的概率再次在进化规则下以多态型均衡时的种群中不同表现型的比例出现了。但现在这个混合均衡是不稳定的。种群比例 x 对平衡点 $x=2/3$ 哪怕是最轻微的偏离都会产生一种累积过程，这个过程将使得种群的混合体进一步远离平衡点。如果 x 从 2/3 处增加，S 类型就变得更具适应性，并且会迅速地繁衍，将 x 增加得更多。倘若 x 从 2/3 处下降，L 类型就变得更具适应性，并且迅速地繁衍，将 x 降低得更多。最终，x 将或者一直上升到 1，或者一直下降到 0，取决于出现什么样的扰动。区别是，在懦夫博弈中，每一种稀有的类型会更具适应性，所以种群比例倾向于做偏离极端位置的运动，并且趋向于某个中间位置的平衡点。相反，在安全博弈中，每一种数量繁多的类型更具适应性；当种群中大多数其余的人都是和你一样的类型时，约会失败的风险就下降了，所以种群比例倾向于向着极端位置运动。

图 13—3 画出了适应性图像及安全博弈的均衡。该图十分类似于图 13—2。这两条线表明了两种类型的适应性与种群比例的关系。两条线的交点给出了平衡点。唯一的区别是，在偏离了平衡点后，数量繁多的类型是更具适应性的，而在图 13—2 中，适应性强的却是数量较少的类型。

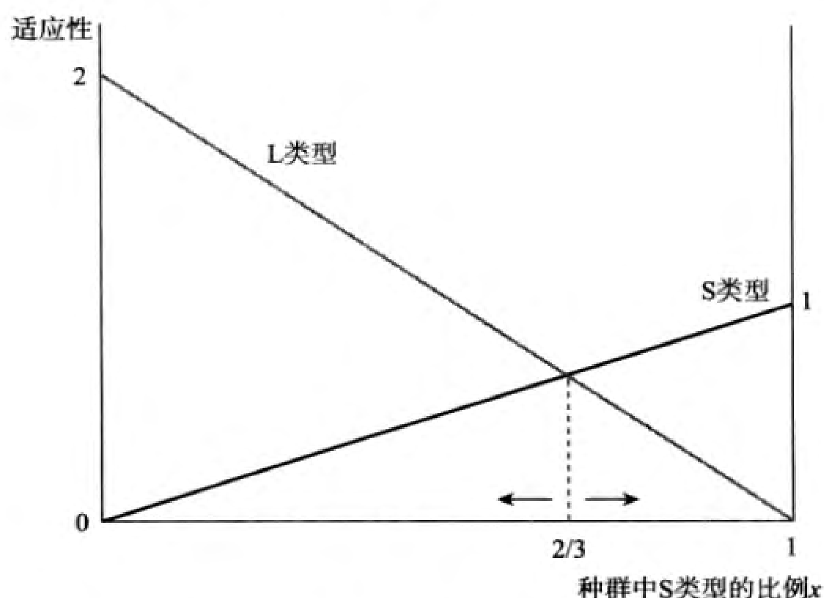


图 13—3　安全博弈的适应性和均衡

因为每一种稀有的类型会不太适应，所以唯有种群的两个极端的单态型结构才是可能的进化稳定结构。不难验证，根据静态检验，两种结果都是 ESS，另一类型的少量变异的入侵将会被消除，因为变异数量的稀少使得它们缺乏适应性。所以，在安全博弈或协调博弈中，与懦夫博弈不同，进化过程并不保留“坏的”均衡。所谓“坏的”均衡，意即参与人选择冲突策略的概率为正。然而，当从任意的初始表现型混合体出发时，动态过程并不能够确保收敛到两个均衡中较好的那一个——种群最后会达到哪一个均衡取决于它是从什么地方开始的。

13.5　物种间的相互作用

要考察的最后一类策略博弈是性别战。在第 4 章里（见表 4—13），我们看到性别战博弈在某些方面与安全博弈看起来是相似的。我们将这两种博弈加以区分的方式是在这里假定“男人”与“女人”仍然对是在星巴克还是本地咖啡店去约会感兴趣——不约会给每人带来一个零支付，但是现在每一种类型喜欢的是不同的咖啡店。因此，为了保证双方采取一致的行动，仍然需要有一定的激励，正如在安全博弈中的那样。但两种可能的双方一致行动的结果是不同的。安全博弈中的不同类型在其偏好上并没有什么不同：两者都喜欢（L，L）胜于喜欢（S，S）。性别战中的参与人在其偏好上是不同的，本地咖啡店给女人带来的支付是 2 而给

男人带来的支付为 1，而星巴克则恰好相反。这些偏好就将两种类型区分开来了。在生物学的语言里，他们就不再能够被看做从由同类组成的动物种群中随机挑选出来的了。[10]其结果就是，他们属于不同的物种（的确如此，恰如男人和女人们相互之间经常认为的那样）。

要从进化的角度研究这种博弈，我们必须将研究方法加以拓展，使其可以运用于从不同物种或不同种群中随机抽取出来的个体间的配对的情形。我们通过将性别战博弈例子加以扩展来说明如何做到这一点。

假定有一个许多男人的种群和一个许多女人的种群。我们从每一种“物种”中随机挑选一个个体，然后要求他们尝试着去约会。所有的男人在去星巴克、本地咖啡店和不约会上的评价（支付）是相同的。类似地，所有的女人也如此。但是，在每一个种群里，某些成员是固执的而其他的却是随和的。固执者总是去他或她所属的“物种”所偏爱的咖啡店。随和者承认其他“物种”有相反的需求因此去别的“物种”愿意去的地方，与其和睦相处。

如果随机抽取恰好抽到一种“物种”的固执者和另外一种“物种”的随和者，那么结果就是固执者所属“物种”所偏好的结果。如果双方都是固执者，就不会有约会。奇怪的是，如果双方都是好说话的，也是这个结果；因为他们都去了对方所偏好的咖啡店。（注意，他们的选择是独立进行的，并且不能够协商。即便他们在事前能够碰面，他们也会出现一种僵局，会说“不，我坚持要满足你的偏好啊”。）

我们修改表 4—13 中的支付表，见表 13—7。之前的可选策略现在被解释为由类型所预先决定的行动（固执的或随和者）。

表 13—7　　　　性别战博弈中的支付

		女人	
		固执者	随和者
男人	固执者	0，0	2，1
	随和者	1，2	0，0

与前面我们所研究过的所有进化博弈相比，这里出现的新特征是行参与人和列参与人是来自不同的“物种”。尽管每种“物种”都是由固执者和随和者所混杂起来的，但两种“物种”内各类型的比例没有理由一定会相同。因此我们必须引入两个变量来表示这两个混合体，并且对两者的动态变化过程进行研究。

设 x 为男人中固执者的比例，而 y 为女人中固执者的比例。考虑一个特定的固执的男人。他会在 y 比例时间里遇到一个固执的女人并且得到一个零支付，以及会在其余的时间里遇到一个随和的女人，并获得 2 个单位的支付。因此，其期望支付（适应性）为 $y\times0+(1-y)\times2=2(1-y)$。类似地，一个随和的男人的适应性为 $y\times1+(1-y)\times0=y$。所以，在男人中，如果 $2(1-y)>y$ 或$y<2/3$，固执者的适应性会更强。固执的男人在他们是更具适应性的时候会迅速增加，即当 $y<2/3$ 时，x 会增加。注意，这一结果的新的、第一眼看起来令人十分惊奇的特征是：在给定“物种”内，每种类型的适应性取决于另一“物种”中各类型

的比例。当然，这也是不足为奇的；记住，现在每一“物种”所进行的博弈都是对抗另一“物种”中的成员的。[11]

类似地，考虑另一“物种”，我们就有这样的结果，即当 $x<2/3$ 时固执的女人是更具适应性的；所以当 $x<2/3$ 时 y 增加。为了直观地理解这个结果，注意到它说的是当另一“物种”没有太多的固执者时，每一“物种”的固执者都要发展得更好一些，因为这时他会更加经常地遇到另一“物种”的随和者。

图 13—4 显示了两“物种”结构的动态变化。x 和 y 都在 0 到 1 之间变化；所以我们有一个单位正方形且 x、y 分别在横轴和纵轴上。其中，垂直线 AB 代表满足 $x=2/3$ 的所有的点，它们都是 y 既不增加也不减少的平衡点。如果当前男人中固执者的比例位于这条线的左边（即 $x<2/3$），y 就增加（将女人中固执者的比例沿着垂直方向向上移动）。如果当前男人中固执者的比例在 AB 的右边（$x>2/3$），那么 y 就下降（垂直向下移动）。类似地，水平线 CD 代表满足 $y=2/3$ 的所有的点，CD 上的所有点都是 x 的平衡点。如果女人中固执者的比例位于这条线的下面（即 $y<2/3$），则男人中固执者的比例 x 就会增加（水平向右）；当 $y>2/3$ 时，这一比例就会下降（水平向左）。

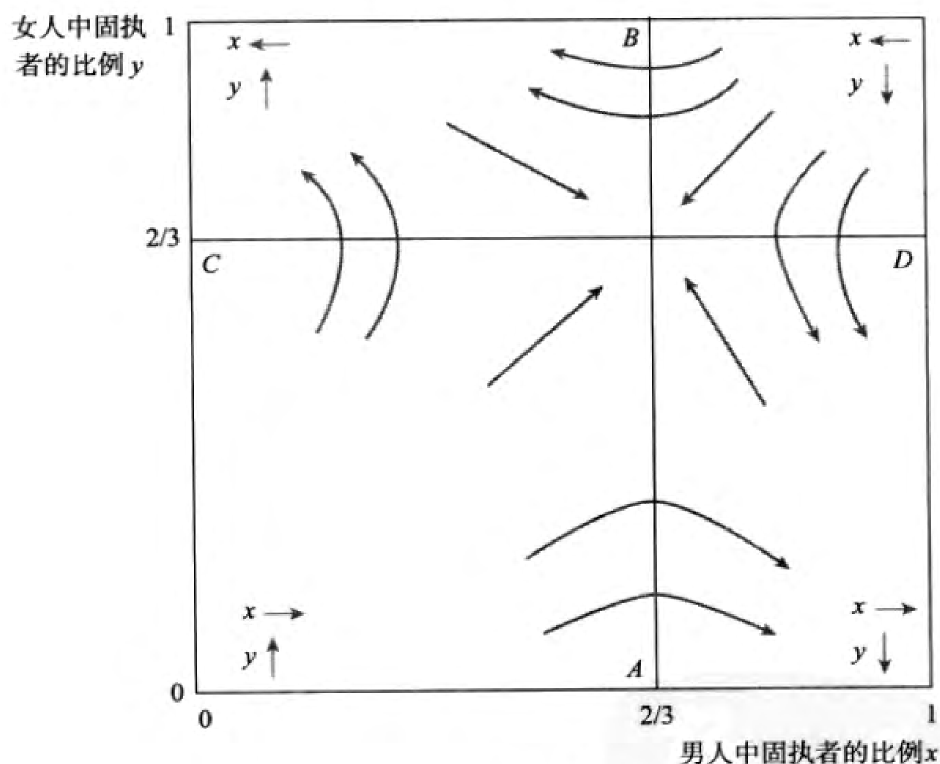


图 13—4　性别战博弈中的“物种”结构的动态变化

当我们将 x 和 y 的变化结合起来考虑时，我们就能够循着它们的动态变化路径去确定种群均衡的位置。例如，从图 13—4 的左下象限中的点出发，动态过程让 x 和 y 都增加。这种联合运动（向着东北方向）会持续到 $x=2/3$ 并且 y 开始下降（现在向着东南方向移动）或者 $y=2/3$ 并且 x 开始下降（现在向着西北方向移动）。在每一个象限中的类似过程都产生出图中所显示的弯曲的动态路径。这些曲线大多朝着图中的东南角或西北角移动，也就是说，它们收敛于（1，0）

或者（0，1）。因此，在大多数情况下，进化动态过程会导致这样的一种结构，其中一种“物种”内全部都是固执者而另外一种“物种”内全是随和者。最终哪一种“物种”包含的是哪种类型则取决于初始条件。要注意的是，如果最初 x 很小并且 y 很大，种群的结构很可能先越过 CD 线，然后向着（0，1）运动——都是固执的女人，$y=1$——而不是先碰到 AB 线然后再向着（1，0）运动；如果初始条件是 x 很小并且 y 很大，我们也可以得到类似的结论。最初有更多的固执者的种群将有终结于全部都是固执者并获得 2 个单位支付的优势。

如果初始比例恰好是平衡的，动态变化将趋向一个多态型点（2/3，2/3）。但是与懦夫博弈中的多态型结果不同，性别战博弈中的多态型是不稳定的。大多数偶然的偏离都会带来一个累积性的运动并最终趋向于两个极端均衡中的某一个：它们就是这个博弈的两个 ESS。这是一个一般性的性质——这样的多物种博弈存在对于每个物种都是单态型的 ESS。

13.6 鹰鸽博弈

鹰鸽博弈（hawk-dove game）是生物学家提出进化博弈理论时所研究的第一个例子。它与对囚徒困境博弈和懦夫博弈到目前为止的分析有很多相同之处，因此这里我们将具体描述这个博弈，以增强和改善对这些概念的理解。

这个博弈并不是指这两种物种之间的博弈，而是指同一物种中两个个体之间的博弈，而鹰和鸽仅仅是对它们策略的命名。背景是对一种资源的竞争。鹰策略表示进攻，即为获得某种数量为 V 的整个资源而战斗。鸽策略表示愿意与别人共享但会逃避战斗。当两个鹰类型相遇时，它们之间就会发生战斗；每一个取胜并且获得 V 的概率均为 1/2，失败受伤得到 $-C$ 的概率也均为 1/2。于是双方的期望支付均为（$V-C$）/2。当是两个鸽类型相遇时，它们共享而非战斗，所以各自得到 $V/2$。如果一个鹰类型与一个鸽类型相遇，后者撤退并得到零支付，而前者得到 V。支付表如表 13—8 所示。

表 13—8　　鹰鸽博弈的支付表

		B	
		鹰	鸽
A	鹰	$(V-C)/2$，$(V-C)/2$	V，0
	鸽	0，V	$V/2$，$V/2$

我们对该博弈的分析类似于囚徒困境博弈和懦夫博弈的情形，所不同的是这里的支付是代数符号。我们将对这个博弈在两种情形下的均衡进行对比：其中一种情形是参与人理性地在鹰策略和鸽策略之间进行选择；另外一种情形是参与人机械地行动并在成功时会获得快速繁殖的奖赏。

13.6.1 理性策略选择与均衡

若$V>C$，则博弈就是囚徒困境博弈，其中鹰策略对应于"背叛"而鸽对应于"合作"。鹰策略是所有参与人的占优策略，但对所有参与人来说，（鸽，鸽）却是一个更好的结果。

若$V<C$，则博弈是一个懦夫博弈。现在有$(V-C)/2<0$，鹰策略不再是一个占优策略。而且，博弈有两个纯策略纳什均衡：（鹰，鸽）和（鸽，鹰）。还存在一个混合策略均衡，其中B选择鹰策略的概率p使得A在鹰策略、鸽策略之间是无差异的，即$p\frac{(V-C)}{2}+(1-p)V=p\times 0+(1-p)\frac{V}{2}$，得到$p=\frac{V}{C}$。

13.6.2 $V>C$时的进化稳定

我们从最初鹰类型占主导的种群开始，并且看看它会不会被一个变异的鸽类型所侵入。我们用d代表变异的鸽类型所占的比例。鹰类型占种群的比例为$(1-d)$。那么，在与随机抽取的对手的对局中，鹰类型将在d时间里遇到鸽类型，并且在每一次这样的场合里获得V单位的支付，在$(1-d)$时间里遇到另外的鹰类型，并且在每一次这样的场合里获得$\frac{(V-C)}{2}$的支付。因此鹰类型的适应性为$\left[dV+(1-d)\frac{(V-C)}{2}\right]$。类似地，一个变异的鸽类型的适应性为$\left[\frac{dV}{2}+(1-d)\cdot 0\right]$。由于$V>C$，故$(V-C)/2>0$。还有$V>0$意味着有$V>V/2$。故对于任何0到之间的$d$值，我们有

$$dV+\frac{(1-d)(V-C)}{2}>d\left(\frac{V}{2}\right)+(1-d)\times 0$$

所以鹰类型更具适应性。鸽类型的变异不能侵入种群。鹰策略是进化稳定的。此种群为单态型种群（全部由鹰类型构成）。

同样的结论对于所有的d值，即任何鸽类型的种群比例来说，都是成立的。因此从任何初始的混合体出发，鹰类型的比例将增加并占据主导。另外，如果种群最初全部是由鸽类型组成的，鹰类型的变异将会侵入并且接管整个种群。因此，动态过程分析证实鹰策略是唯一的ESS。这种代数分析证明并一般化了我们先前在餐馆定价囚徒困境博弈数值例子中的发现。

13.6.3 $V<C$时的进化稳定

如果最初种群中鹰类型占主导，其中有一个很小比例d的鸽类型变异，则双

方都有同 13.6.2 节一样的适应性函数。然而，当$V<C$时，$(V-C)/2<0$。我们仍然有$V>0$，故 $V>V/2$。但是由于 d 很小，含有（$1-d$）的比较项比起含有 d 的比较项来说是远为重要的，故

$$d\left(\frac{V}{2}\right)+(1-d)\times 0>dV+\frac{(1-d)(V-C)}{2}$$

因此鸽类型变异比占主导的鹰类型有较强的适应性，并且能成功地侵入种群。

但如果最初种群中几乎全部是鸽类型，则我们必须考虑是否一个小的比例 h 的鹰类型变异能否成功侵入种群（注意，由于这里变异的是鹰类型，我们用 h 表示变异入侵者的比例）。鹰类型变异的适应性为$[h(V-C)/2+(1-h)V$，与之对照的是鸽类型的适应性为$[h\times 0+(1-h)(V/2)]$。再一次，$V<C$ 意味着 $(V-C)/2<0$，且 $V>0$ 意味着 $V>V/2$。但是当 h 很小时，我们有

$$\frac{h(V-C)}{2}+(1-h)V>h\times 0+(1-h)\left(\frac{V}{2}\right)$$

该不等式表明鹰类型有较强的适应性，并且将能成功地侵入一个鸽类型的种群。所以每一种类型的变异都能成功侵入其他类型的种群。种群不可能是单态型的，且任何纯的表现型都不可能是一种 ESS。代数分析又证实了我们在前面的懦夫博弈例子中所获得的发现（表 13—5 和图 13—2）。

当 $V<C$ 时种群会发生什么呢？有两种可能性存在。一种是每个参与人都遵循一种纯策略，但种群却是一个由遵循不同策略的参与人组成的一个稳定的混合体。这恰好是我们在 13.3 节里通过对懦夫博弈的分析所得到的多态型均衡。另一种可能性就是每个参与人采用一种混合策略。我们从多态型的情形开始。

13.6.4 $V<C$：稳定多态型种群

当种群中鹰类型的比例为 h 时，鹰类型的适应性为 $h(V-C)/2+(1-h)V$，而鸽类型的适应性为 $h\times 0+(1-h)(V/2)$。鹰类型更具适应性，如果

$$h(V-C)/2+(1-h)V>(1-h)(V/2)$$

将其进行化简得

$$h(V-C)/2+(1-h)V/2>0$$
$$V-hC>0$$
$$h<V/C$$

当 $h>V/C$ 或 $1-h<1-V/C=(C-V)/C$ 时，鸽类型的适应性就更强。可见，每一种更加稀少的类型更具适应性。因此我们在平衡点得到一个稳定多态型的均衡，其中种群中鹰类型的比例是 $h=V/C$。这恰好等于我们在 13.6.1 节中曾计算过的理性行为假定下博弈的混合策略纳什均衡中每个个体选择鹰策略的概率。我们又一次从进化的角度对懦夫博弈中的混合策略结果给出了一个理由。

我们把将这种情形画一个类似于图 13—2 的图形的任务留给读者自己去完

成。这需要读者首先弄清楚每种类型的种群比例收敛于稳定均衡混合体的动态变化过程。

13.6.5 $V<C$：每个参与人都采用混合策略

回忆 13.6.1 节中计算出的理性决策博弈的均衡混合策略，其中 $p=V/C$ 是选择鹰策略的概率，而 $(1-p)$ 是选择鸽策略的概率。是否存在着进化版本的对应者，其中表现型用的是混合策略呢？让我们来考察一下这种可能性。我们仍然有采用纯鹰策略的 H 类型和采用纯鸽策略的 D 类型，但现在被称为 M 类型的第三种表现型可能存在，这样一种类型采用的是一种混合策略，其中鹰策略的概率为 $p=V/C$，鸽策略的概率为 $1-p=1-V/C=(C-V)/C$。

若 H 类型或 D 类型遇到 M 类型，他们的期望支付取决于 p，即 M 类型选择鹰策略的概率，以及 $(1-p)$，即 M 类型选择鸽策略的概率。那么，每个参与人得到的支付为他对付 H 类型时的支付的 p 倍，加上他对付 D 类型时的支付的 $(1-p)$ 倍。所以当 H 类型遇到 M 类型时，其期望支付为

$$
\begin{aligned}
p\frac{V-C}{2}+(1-p)V &= \frac{V}{C}\frac{V-C}{2}-\frac{C-V}{C}V \\
&= -\frac{1}{2}\frac{V}{C}(C-V)+\frac{V}{C}(C-V) \\
&= V\frac{(C-V)}{2C}
\end{aligned}
$$

并且当 D 类型遇到 M 类型时，他得到

$$
p\times 0+(1-p)\frac{V}{2}=\frac{C-V}{C}\frac{V}{2}=\frac{V(C-V)}{2C}
$$

这两个适应性是相同的。这是不足为奇的；混合策略的比例就是通过这个等式来决定的。因此一个 M 类型遇到另一个 M 类型也得到同样的期望支付。为了后面在采用时方便简捷，我们把这个共同的支付称为 K，其中 $K=\frac{V(C-V)}{2C}$。

当我们从进化稳定性角度来考察 M 类型时，这些等式就凸显出一个问题。假设某种群完全由 M 类型组成并且有少量的 H 类型变异侵入，二者组成的总的种群中 H 类型占一个很小的比例 h，则典型的变异得到的期望支付为

$$
h(V-C)/2+(1-h)K
$$

为了计算 M 类型的期望支付，注意，他会在相互作用的 $(1-h)$ 比例时间里遇到另一个 M 类型，并且每一次都获得支付 K。于是他会在 h 比例时间里遇到 H 类型；在这些场合他在 p 比例时间里采用鹰策略并得到 $(V-C)/2$，在 $(1-p)$ 比例时间里采用鸽策略并得到 0。因此其总的期望支付（适应性）就为

$$
hp(V-C)/2+(1-h)K
$$

由于 h 很小，M 类型与 H 类型的适应性几乎是相等的。要点是当只有很少的变异时，H 类型和 M 类型在大多数时间里都只能遇到 M，所以在这种互动中恰如我们看到的那样两者有相同的适应性。

因此进化稳定性依赖于这样一种差别，即最初的种群中的 M 类型和 H 类型变异各自在与一种少量的变异对局时，前者的适应性是否强于后者。从代数上看，当 $pV(C-V)/2C=pK>(V-C)/2$ 时①，M 类型比起 H 类型来说在与其他 H 类型的变异相遇时要有更强的适应性。在我们这里的例子中，这个条件是成立的，因为 $V<C$（所以 $V-C$ 是负的）且 K 是正的。直观地看，该条件告诉我们一个 H 类型变异在应对另外一个 H 类型变异时总是比较糟糕的，这是由于战斗会付出很高的成本；而 M 类型只有一部分时间在战斗，且因此只在 p 比例时间里付出这种成本。总的看来，当与变异互动时，M 类型要表现得出色一些。

类似地，一种鸽类型的变异能否成功侵入 M 类型种群也取决于鸽类型变异的适应性与 M 类型的适应性之间的大小比较。如前所述，变异有 d 比例的时间遇到另外的 D 类型，在 $(1-d)$ 比例的时间里面临 M 类型。一个 M 类型也在 $(1-d)$ 比例的时间里遇到另外一个 M 类型。但是在 d 比例的时间里，M 类型面临一个 D 类型，并且在这些时间内的 p 比例里采用鹰策略，因而得到 pV；在这些时间的 $(1-p)$ 的比例里采用鸽策略，因而得到 $(1-p)V/2$。鸽类型的适应性因此是 $[dV/2+(1-d)K]$，而 M 类型的适应性就是 $d\times[pV+(1-p)V/2]+(1-d)K$。在每一个适应性表达式中，最后一项都是相同的，所以仅当 $V/2$ 大于 $pV+(1-p)V/2$ 时，鸽类型的入侵才是成功的。这个条件是不成立的；因为后面的表达式包括 V 和 $V/2$ 的一个加权平均，而且一旦有 $V>0$，它必定大于 $V/2$。因此，鸽类型的入侵也不会成功。

这样的分析告诉我们，M 策略是一个 ESS。因此若有 $V<C$，种群就可以表现出两种进化稳定结果中的任何一种。一种稳定结果是种群中存在不同的类型（一种稳定的多态型）；另一种稳定结果是种群中只有一种类型，此类型以与多态型中不同类型的比例相同的概率选择纯策略。

13.7 种群中含有三种表现型

如果种群中只有两种可能的表现型（策略），我们就可以通过将所考虑的类型与唯一的变异类型进行比较来完成 ESS 的静态检验，并可用类似于图 13—1、图 13—2 和图 13—3 的图形来展现种群在一个进化博弈中的动态变化过程。现在，我们来说明如果种群中有三种（甚至更多种的）可能的表现型的话，这种思路和方法如何能够加以利用，以及会出现什么样的新的考虑。

① 疑原书有误，M 类型与 H 类型相遇时其支付为 $p(C-V)/2$。——译者注

13.7.1 检验 ESS

让我们通过引入第三种可能的表现型来重新考察 13.2.1 节和表 13—3 中的三次重复囚徒困境博弈。这种策略，用 N 表示，即绝不背叛。表 13—9 给出了三种策略 A，T 和 N 的适应性表。

为了检验任何一种这样的策略是否就是一个 ESS，我们考虑一个全部由一种类型组成的种群是否能够被其他类型的变异所侵入。比如，一个全部由 A 类型组成的种群，是不会被 T 类型或 N 类型的变异所侵入的，因此 A 类型是一个 ESS。而全部由 N 类型组成的种群会被 A 类型的变异侵入，所以 N 类型不是一个 ESS。T 类型又怎么样呢？全部由 T 类型组成的种群不会被 A 类型的变异所侵入，但是它可以被 N 类型的变异所侵入，所以 T 类型也不是一个 ESS。综上所述，只存在一个单态型 ESS，即策略 A。

表 13—9　　三种类型的三次重复囚徒困境博弈（百美元/月）

		列		
		A	T	N
行	A	864，864	936，792	1 080，648
	T	792，936	972，972	972，972
	N	648，1 080	972，972	972，972

让我们来进一步考虑一个全部由 T 类型组成的种群被 N 类型的变异所侵入的情形。这两种类型可以暂时快乐地共存一段时间。但是，倘若该种群中 N 类型所占的比例过大，则 A 类型变异就能够入侵；A 类型在应对 N 类型时表现较好，但在面对 T 类型时表现就比较糟糕。我们可以具体考虑这样一个种群，其中 N 类型的比例为 x 而 T 类型的比例为 $(1-x)$。每一种类型的适应性皆为 972。在这个种群中 A 类型变异的适应性是 $936(1-x)+1\,080x=144x+936$。如果有 $144x>972-936=36$ 或 $x>1/4$，它就超过 972。因此，只要 N 类型的比例小于 25%，我们就能够有另外一个由 T 类型和 N 类型混合而成的多态型均衡。

13.7.2 动态分析

为了展开对有三种可能表现型博弈的动态分析，我们转向另一著名博弈，即石头—剪刀—布（RPS）。在这种博弈的理性决策博弈理论中，每一个参与人都在三种可选择的行动中选择一种，或是石头（出拳头），或是布（张开手掌），或是剪刀（用两根手指做出剪刀的模样）。该博弈的规则是：石头击败（“砸碎”）剪刀，剪刀击败（“剪破”）布，以及布击败（“包住”）石头；同样的行动就打个

平手。倘若参与人选择的是不同的行动，赢家就得到支付 1 而输家获得支付 −1；平局使得两个参与人都得到 0。

关于进化的例子，我们讨论生活在加利福尼亚海岸的单面斑点蜥蜴（side-blotched lizard）所面临的情境。这一物种存在三种类型的雄性交配行为，每一种类型都有一种特定的颈前部颜色。蓝色颈前部的雄性保护着一小群雌性伴侣并抵挡来自那些偷偷地溜进来想与未得到保护的雌性进行交配的黄色颈前部的雄性的求爱。黄色颈前部雄性的溜进策略在与那些拥有许多配偶且经常出去进攻性地追逐更多伴侣的橙色颈前部雄性的争斗中会取胜。雌性单面斑点蜥蜴倾向于选择那些可以被橙色颈前部雄性的入侵所征服的蓝色颈前部的雄性。[12]它们之间的相互作用可以用表 13—10 中的 RPS 博弈的支付结构来描述。表中，我们设立了一列 q—mix，并由此考察该博弈混合策略均衡的进化等价（种群中各种类型按一定比例组合）。[13]

表 13—10　　　　三种类型进化博弈的支付

		列			
		黄色颈前部	蓝色颈前部	橙色颈前部	q—mix
行	黄色颈前部	0	−1	1	$-q_2+(1-q_1-q_2)$
	蓝色颈前部	1	0	−1	$q_1-(1-q_1-q_2)$
	橙色颈前部	−1	1	0	$-q_1+q_2$

设 q_1 为种群中黄色颈前部蜥蜴的比例，q_2 为蓝色颈前部蜥蜴的比例，那么 $(1-q_1-q_2)$ 就是橙色颈前部蜥蜴的比例。支付表右边的一列为每个行参与人在与这个表现型的混合体相遇时的支付，也就是说，正好是行的适应性。恰如在单面斑点蜥蜴种群中已经看到的那样，假定适应性为正的类型在种群中的比例会增加，而适应性为负的类型在种群中的比例会减少[14]，那么，

$$q_1 \text{增加当且仅当} -q_2+(1-q_1-q_2)>0 \text{ 或 } q_1+2q_2<1$$

当 q_2 即蓝色颈前部类型的比例较小时，或者 $(1-q_1-q_2)$ 即橙色颈前部类型的比例较大时，种群中黄色颈前部类型的比例就会增加。这是合理的；因为黄色颈前部蜥蜴在与蓝色颈前部蜥蜴对局时会表现得很糟糕，而与橙色颈前部蜥蜴对局时会表现得很不错。类似地，我们知道

$$q_2 \text{增加当且仅当 } q_1-(1-q_1-q_2)>0 \text{ 或 } 2q_1+q_2>1$$

当黄色颈前部竞争者的比例较大或橙色颈前部类型的比例较小时，蓝色颈前部雄性就有好日子过了。图 13—5 用几何图形显示出了该博弈的种群动态特征及所得到的均衡。图中的三角形区域由两轴及直线 $q_1+q_2=1$ 定义，它包含了所有可能的 q_1 和 q_2 的均衡组合。三角形内还有两条直线。第一条是 $q_1+2q_2=1$（较为平缓

的那一条），它是 q_1 的平衡线。对于这条直线以下的 q_1 和 q_2 的组合，q_1（黄色颈前部类型的比例）增加；而对于在这条直线以上的组合，q_1 减少。第二条陡峭的直线是 $2q_1+q_2=1$，它是 q_2 的平衡线。在这条直线的右边（$2q_1+q_2>1$），q_2 增加，而直线的左边（$2q_1+q_2<1$），q_2 减少。图中的箭头表明的是这些种群比例的变动方向，标有箭头的曲线则显示了典型的动态变化路径。一般性的思想与表 13—9 一样。

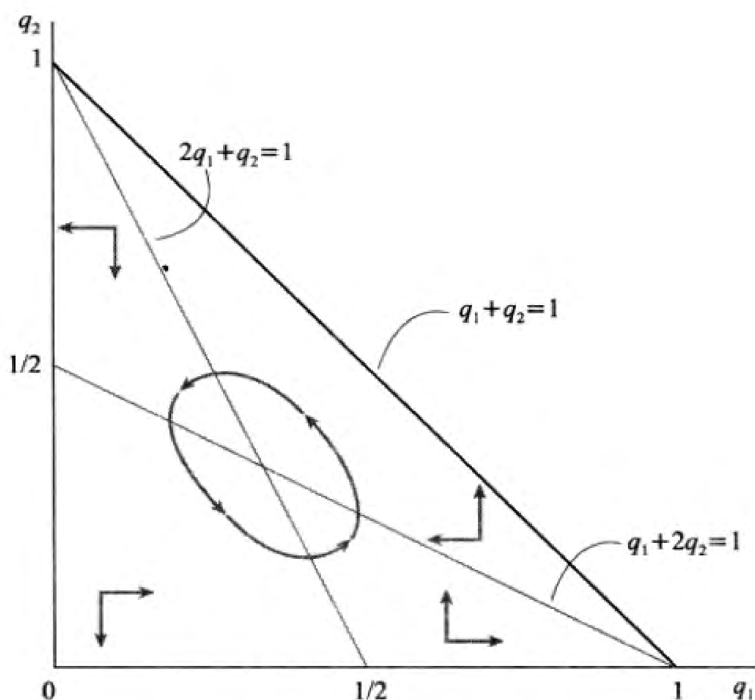


图 13—5　进化 RPS 博弈的种群动态变化

图中的两条直线都包括了 q_1 和 q_2 既不上升也不下降的点。因此，两条直线的交点代表了 q_1、q_2 从而（$1-q_1-q_2$）都不变的点；这个点对应于一个多态型均衡。不难验证，这里有 $q_1=q_2=1-q_1-q_2=1/3$。这一比例与 RPS 博弈中的理性决策混合策略均衡中的概率是相同的。

这个多态型结果是稳定的吗？我们还不能肯定。动态变化路径是以该点为中心的一个圈（图 13—5 表明它是一个椭圆）。这条路径是绕着交点螺旋式地收缩（此时我们得到稳定的结果）还是螺旋式地从中心点向外扩展（此时该点就是不稳定的），取决于种群比例对于适应性的精确反应。所画的路径圈还甚至可能既不趋近也不远离这个均衡点。

有关证据指出单面斑点蜥蜴种群结构是以各种类型等比例共存的那一点（即多态型均衡点）为圆心绕圈变化的，其中一种类型在一个为期若干年的周期里会稍微更常见，但随后就会被更强的竞争者所超越。这样的循环是否趋于稳定的均衡仍然是一个需要在未来进行研究的课题。在进化博弈中，另外一个 RPS 类型相互作用的例子涉及与大肠杆菌（E. coli bacteria）相关的三种食物中毒系。每一种都取代另一种但是又被第三种所取代，正如前面在三种类型博弈中所描述的那样。研究这三种系之间相互作用的科学家已经证明，倘若每两种系之间的相互作

用是局部的，并且每种系以小团块的形式连续地改变其位置，则一个多态型均衡就能够得以维持。[15]

13.8 部分一般性理论

为了获得可以进一步应用的理论框架和研究方法，我们现在把前面 13.6 节里所介绍的思路一般化。这就不可避免地需要运用一些稍微抽象一点的概念和一点点代数知识。因此，我们仅仅考虑单一物种内单态型均衡的情形。数学功底好的读者可以将此方法推广到两种物种的多态型情形。而对本节内容不感兴趣或者没有相应准备的读者可以跳过本节而不会失去连续性。[16]

我们来考虑某个单一物种内随机配对，其中可供选择的策略有 I，J，K，…，其中一些可以是纯策略，另一些可以是混合策略。每一个个体成员都天生地只选择其中某个策略。我们设 $E(\mathrm{I},\mathrm{J})$ 表示 I 类型参与人在遇到 J 类型参与人时所获得的单阶段支付。那么 I 类型参与人遇到同类型参与人的单阶段支付就是 $E(\mathrm{I},\mathrm{I})$。我们记 $W(\mathrm{I})$ 为 I 类型参与人的适应性。当遇到某类参与人的概率就恰好等于这类参与人在种群中的比例时，这也就是他随机地遇到其他对手时的期望支付。

假设种群中全是 I 类型的个体。我们来考察这种种群结构是否是进化稳定的。为此，我们想象一下种群被少量 J 类型变异侵入的情形，假定种群中变异的比例是一个非常小的数 m。现在 I 类型的适应性为

$$W(\mathrm{I})=mE(\mathrm{I},\mathrm{J})+(1-m)E(\mathrm{I},\mathrm{I})$$

变异类型的适应性为

$$W(\mathrm{J})=mE(\mathrm{J},\mathrm{J})+(1-m)E(\mathrm{J},\mathrm{I})$$

因此，种群中主要类型与变异类型的适应性之差为

$$\begin{aligned}W(\mathrm{I})-W(\mathrm{J})=&m[E(\mathrm{I},\mathrm{J})-E(\mathrm{J},\mathrm{J})]\\&+(1-m)[E(\mathrm{I},\mathrm{I})-E(\mathrm{J},\mathrm{I})]\end{aligned}$$

由于 m 很小，如果前面表达式中的第二部分为正的话，主要类型的适应性将会比变异类型的要强，即有

$$W(\mathrm{I})>W(\mathrm{J}) \text{ 在 } E(\mathrm{I},\mathrm{I})>E(\mathrm{J},\mathrm{I}) \text{ 时成立}$$

所以，种群中的主要类型是不可能被侵入的；当与主要类型中的一个成员配对时，主要类型要比变异类型更具适应性。这就形成了进化稳定的**原生准则**（primary criterion）。相反，当 $E(\mathrm{I},\mathrm{I})<E(\mathrm{J},\mathrm{I})$ 时，$W(\mathrm{I})<W(\mathrm{J})$，J 类型变异就能够成功侵入，并且完全由 I 类型组成的种群不会是进化稳定的。

然而，可能有 $E(\mathrm{I},\mathrm{I})=E(\mathrm{J},\mathrm{I})$。如果种群最初由采用将 I 策略与 J 策略混合的策略的单个表现型组成（一个混合策略的单态型均衡），就的确会出现这种

情况，就像鹰鸽博弈中的最后一种情形那样（见 13.6.5 节）。这时 $W(\mathrm{I})$ 与 $W(\mathrm{J})$ 的差就由每种类型面对变异类型时的支付所决定了。[17] 当 $E(\mathrm{I},\mathrm{I})=E(\mathrm{J},\mathrm{I})$ 时，如果 $E(\mathrm{I},\mathrm{J})>E(\mathrm{J},\mathrm{J})$，我们得到 $W(\mathrm{I})>W(\mathrm{J})$。这是判断策略 I 的进化稳定性的**次生准则**（secondary criterion）。只有当原生准则不能给出结果时，即只有当 $E(\mathrm{I},\mathrm{J})=E(\mathrm{J},\mathrm{I})$ 时，才运用这个准则。

原生准则是十分重要的。它说的是，如果策略 I 是进化稳定的，那么对于变异可能选择的所有其他的策略 J，都有 $E(\mathrm{I},\mathrm{I})>E(\mathrm{J},\mathrm{I})$。这意味着策略 I 是对于自身的最优反应。换句话说，如果这个种群中的个体突然变成了理性计算者，那么每个个体都选择策略 I 将会是一个纳什均衡。因此进化稳定就意味着对应的理性决策博弈的纳什均衡！

这是一个非常值得注意的结果。如果你对前几章纳什均衡理论中潜在的理性行为假设不满意，并且你试图在进化博弈论中去寻找更好的解释，你将发现它带来的是相同的结论。这一非常有吸引力的生物学描述——提出非最大化行为的假设，但根据适应性进行选择——并未产生任何新的东西。即便是有，那也只是为纳什均衡提供了一个支撑性的辩护而已。当一个博弈有几个纳什均衡的时候，进化博弈甚至会提供一个很好的理由去从中挑选出一个合理的均衡。

但是，你在对纳什均衡的信心得到增强的同时也要悠着点儿。我们关于进化稳定的定义是静态的而不是动态的。它只考察了我们对之进行检验以确定其是否是均衡的种群结构（单态型或者按一定的正确比例组成的多态型）会不会被一小部分变异成功侵入，而没有考察从最初任意的类型组合开始，所有不需要的类型是否会衰亡以及种群能否达到均衡结构的状态。考察的对象也只包含逻辑上可能的变异类型；如果理论家们没有进行正确的分类并且被忽略了的某些类型的变异实际上会产生，这种变异类型就会成功地侵入并摧毁假定的均衡。我们在 13.2.1 节分析两次重复的囚徒困境博弈时最后所做的评论中，曾警告了这种可能性，并且习题 6 表明了它会怎么样产生。最后，在 13.7 节中我们也曾指出过进化的动态过程可能会因某些原因而根本不会收敛。

13.9 领地博弈

我们到目前为止考察了这样的一些情形，其中每一个博弈都是正好在从种群中随机性地挑选出来的两个参与人之间进行的。然而，当整个种群进行博弈时，就存在着另外的情形。在生物学里，一群具有遗传性行为的动物可能为了某种资源或领地而竞争。在经济或商业领域中，每个行业里有许多企业，每一个都遵循着由其企业文化所决定的策略，所有企业都会互相竞争。

这种进化博弈与第 12 章中的理性决策集体行动博弈的关系，和前几节里的逐对进行的进化博弈与第 4 章到第 8 章的理性决策两人博弈的关系是相同的。正如我们将这些章节里的期望支付表转换为图 13—1 和图 13—2 以及图 13—3 中的适应性图形一样，我们能够将集体行动博弈图形（图 12—1 到图 12—3）

转换为进化博弈的适应性图形。例如，考虑一种动物物种，其所有成员生活在一个共同的饲养场里。存在着两种表现型：一种进攻性地为了食物而战斗；另外的一种围绕着食物走来走去，并且在可能的情况下偷偷地吃一些。如果进攻性的那一种的比例很小，它们就生存得很好，但是，倘若它们的数量太多的话，偷吃者通过漠视正在进行的战斗而生存得好。这是一种其适应性图形完全类似于图 12—2 的集体懦夫博弈。因为不需要新的原理和技术，我们将其留给读者自己去进一步探讨。

13.10 合作及利他的进化

进化博弈论有两个基本思想：首先，单个生物是与同一物种中的其他生物或其他物种中的成员进行博弈的；其次，带来高支付（更具适应性）的策略的遗传型会不断增加，而其他的遗传型在种群中的比例则会不断下降。这些思想令人联想到了某些达尔文的诠释者们所说的那种为了生存而激烈斗争的原则，他们按照字面上的意思去理解“适者生存”，将“自然存在于牙齿与爪子之中”的图景呈现于人们脑中。但事实上，自然界里也表现出许许多多合作的（其中动物个体的行为为种群中的其他每个个体带来很大的利益），甚至是利他主义的（其中动物个体牺牲自己的较大利益去给别的个体带来好处）行为。在动物界中，蜜蜂、蚂蚁就是最明显的例子。这样的行为能不能与进化博弈论的观点相协调呢?

生物学家们运用四重分类的方法对自利性动物（或者表现型或者基因）中可能出现的合作和利他的方式进行了分类。李・杜盖金（Lee Dugatkin）命名了四种种类：（1）家族动力学；（2）互惠性交易；（3）自利性团队工作；（4）群体利他主义。[18]

作为家族动力学的一个例子，蚂蚁和蜜蜂的行为大概是最容易理解的了。蚁群或蜂群中的所有个体是密切相关的，因此在很大程度上它们拥有相同的基因。蚁群中所有工蚁都是姊妹，它们有一半的基因相同；因此，从基因遗传的意义上来说，另外两只工蚁的生存与自己的生存同等重要。同一蜂房中所有工蜂之间都是半个姊妹的关系，它们有四分之一的基因相同。当然，单个蚂蚁或蜜蜂并不能计算出牺牲自己去保全两只同类的蚂蚁或四只同类的蜜蜂是否值得，但其成员有着这种行为（表现型）的群体的潜在基因将因此而繁衍不息。“进化最终是运行于基因水平上的进化”这一观点在生物学上有着巨大影响和深远的意义，尽管它曾经被许多人所滥用，就像达尔文“自然选择”的原初思想被人们所滥用了一样。[19]有趣的是“自私的基因”可以通过其在一个更大的基因组织如细胞中表现出无私的行为而得以成功地存续。类似地，一个细胞和它的基因也可通过参与合作和接受它们在一个组织体内的指定任务而成功地存续下去。

互惠性的利他主义可以出现在相同的或不同的物种里的没有关系的个体成员之间。这种行为实质上是重复囚徒困境解的一个例子，其中参与人采用类似以牙还牙的策略。例如，一些以寄生虫为食的鱼虾常到大鱼的嘴里或腮中觅食；大鱼

也由于这种“清洁服务”而让它们在自己的嘴里自由游动而不伤害它们。一个更加有吸引力——虽然有些恐怖——的例子是吸血蝙蝠会与那些没有找到血的蝙蝠共同享用血液。在一个实验中，从不同的地方找来的蝙蝠被放在一起并且被挑选出来挨饿，只有那些即将饿死的蝙蝠（在24小时内不进食就会死亡）被实验中的其他蝙蝠给予了血液。然而关键的一点是，蝙蝠只是从它们的“老乡”那里得到血液，并且吸血蝙蝠更多地是将已咽下的血液吐出来再给予那些求助的个体。[20]再次申明，我们并没有假设每只蝙蝠有意识地计算过是“继续合作”（给予血液）还是“背叛”（不给予血液）对自己更有利。相反，这一行为出于本能。

当所有其他个体都选择合作时，选择合作是符合每一个个体的利益的，此时，自利性团队工作就会出现。换句话说，这类合作行为适用于挑选出安全博弈中的好的结果。杜盖金指出，种群在严酷的环境里倾向于比在不太严酷的环境中更加容易出现自利性团队工作。当条件很糟糕时，群体中任何一个个体因为偷懒而逃避工作都会给包括偷懒者自己在内的整个群体带来灾难。于是在这样的条件下，每个个体对于生存来说都是至关重要的，并且只要其他个体都在努力工作的话就没有个体会因为偷懒而逃避工作。在较好的环境里，每个个体都会希望成为其他个体努力的搭便车者，而这样的行为并不威胁到包括它自己在内的整个群体之生存。

接下来我们走出生物学到社会学中去：一个有机体（和它的细胞，最终它的基因）将获益于其在一个有机体集合内的合作性表现，这个有机体集合就是社会。这带给我们一种群体利他主义的思想，说明即使在个体之间并非近亲关系的种群中我们也应可看到某种合作行为。我们的确找到了这种例子。像狼那样的食肉动物群就是一个很好的例子，而大猩猩的群落常常表现得像一个大家庭。即使在作为被捕食者的种群中，在需要个体去轮流监视食肉猛兽的行踪时也会出现合作行为。不同物种间也会有合作。

一般说来，一个其中的成员表现出合作行为的群体在与其他的群体进行相互作用时，要比其中的成员在群体中寻求搭便车的群体更加成功一些。倘若在进化动态演化的特定框架中，群体之间的选择是比群体内部的选择更加强烈的一种力量的话，则我们就会看到群体性利他主义。[21]

个体的本能是通过基因固定在个体大脑中的，但互利和合作也可能出现在种群内有目的的思考或尝试中，并通过社会化而传播开——这种传播不是通过基因，而是通过明确的指导和对年长者行为的观察来实现的。这两种渠道——养育和本性——的相对重要性因物种和环境的不同而不同。你可能认为社会化对于人类来说是更为重要的，但实际上，它对其他动物也非常重要。这里我们引用一个非同寻常的例子。1911—1912年，罗伯特·F·斯科特（Robert F. Scott）到南极探险时，使用了一群西伯利亚狗。这群狗是一起长大的并为了南极探险而接受过专门训练。在短短的几个月里，它们就建立了一种合作机制，靠惩罚来维持合作。“它们奋力合作并用巨大的努力去惩罚那些未尽力拉或者过分用力拉的同伴……它们惩罚的方法总是相同并且有效的，在它们看来是公平的惩罚在我们看来却是谋杀。”[22]

这是对合作行为如何与进化博弈论相一致的一个令人鼓舞的说明，并且也可由此看出自私行为的困境是可以被克服的。事实上，研究利他行为的科学家们最近报告了在人类中存在利他性惩罚或强烈的互惠性（它区别于互惠性利他主义）的实验证据。他们的证据表明人们愿意去惩罚那些在群体中不努力的家伙，即使这样做代价高昂且不存在预期收益。如果具有这种特征的群体能够更好地在战争与灾难性事件中幸存下来，这种强烈的互惠性倾向将有助于解释人类文明的兴起。[23]尽管有这些发现，强烈的互惠性并未在动物世界中广泛传播开来。"与解释了蚂蚁之间和会照顾其幼子的生物之间的合作的裙带关系相比较，互利行为被证明是稀缺的。这大概是由于这样一个事实，即互利行为不仅需要重复性的互动，而且还需要识别及记住其他个体的能力。"[24]换句话说，在11.2.4节的理论分析中得到的解决重复性囚徒困境的必要条件看来在进化博弈中是起作用的。

13.11 总　　结

生物学的进化论类似于社会科学家所使用的博弈论。进化博弈是在由基因事先决定的不同行为表现型之间而不是理性选择的策略之间进行的。在进化博弈中，具有较强适应性的表现型生存下来，在与其他表现型的重复性互动中不断繁殖，并在种群中增加它们的比例。一个由一种或者多种表现型以一定比例组成的种群被称为是进化稳定的，如果它不能被其他的变异表现型成功入侵，或者如果它是更具适应性的表现型繁殖动态过程的极限结果的话。倘若一种表现型在面临一种变异表现型的侵入时，仍能保持自己的主导性，我们就称这种表现型是进化稳定策略，而且称单独由它所组成的种群表现出单态型。如果两种或更多种表现型共同存在于一个进化稳定的种群，它就被称为是表现出多态型的。

当把进化博弈理论更一般地用于非生物博弈时，单个参与人所遵循的策略就被理解为经验法则，而不是由基因决定的。繁殖过程代表着更一般的传播方式，包括社会化、教育以及模仿；变异代表着新策略的尝试。

进化博弈具有类似于第4章和第8章中所分析过的支付结构，包括囚徒困境和懦夫博弈。在每一种情形里，进化稳定策略折射出来的或者是由理性决策参与人所进行的同样结构的博弈的纯策略纳什均衡，或者是在这样一个博弈中的混合策略均衡中的比例。在囚徒困境中，"总是背叛"是进化稳定的；在懦夫博弈中，稀有的类型更具适应性，故存在着一种多态型的均衡；在安全博弈中，稀有的类型适应性较差，故多态型的结构是不稳定的，且均衡位于极端的位置。当博弈在两种不同物种的每一种中的两个不同类型成员之间进行时，可运用更为复杂但类似的结构分析去求解均衡。

鹰鸽博弈是经典的生物学例子。对此博弈的分析与对进化博弈的囚徒困境及懦夫博弈版本的分析相似；进化稳定策略依赖于支付结构的特征。当有两种以上的类型相互作用时或在非常一般性的情况下，也能够进行分析。该理论表

明进化稳定的要求产生出一种均衡策略，它等价于理性决策参与人所得到的纳什均衡。

关键术语

进化稳定（evolution stable）
进化稳定策略（evolutionary stable strategy, ESS）
适应性（fitness）
遗传型（genotype）
鹰鸽博弈（hawk-dove game）
相互作用（interaction）
变异侵入（invasion by a mutant）
单态型（monomorphism）
变异（mutation）
表现型（phenotype）
领地博弈（playing the field）
多态型（polymorphism）
原生准则（primary criterion）
次生准则（secondary criterion）
选择（selection）

习　题

1. 证明下面的陈述："如果在理性参与人进行的博弈的支付表中，某策略是被严格占优的，则在同一博弈的进化版本中，无论最初的种群混合体如何，它最终都会消亡。如果该策略被弱占优，则它可以与某些其他类型共存，但不会存在于有所有类型的混合体中。"
2. 考虑一个有许多参与人的大型种群，其中存在两种可能的表现型。这些参与人相遇并且进行一场生存博弈，他们或者是进行战斗或者是共享同一食物源。一种表现型总是战斗，另一种总是共享；对于这个问题，假定在种群中没有其他的变异会出现。假设食物源的价值为 200 卡路里，且卡路里的吸收决定了每个参与人的繁殖适应性。如果两种共享型相遇，则每一个都获得一半的食物。但是，倘若一个共享型与一个战斗型相遇，则共享型会立即将食物让给战斗型，战斗型获得所有的食物。
   (a) 假设一次战斗的成本是 50 卡路里（对于每一个战斗型来说）且在两个战斗型相遇时，每一个都有同样的可能性（概率 1/2）赢得战斗和食物或失败且得不到食物。画出从这个种群里随机挑选出来的两个参与人之间的博弈的支付表。找出该种群中的所有 ESS。此时进行的是什么类型的博弈？
   (b) 现在假设一次战斗的成本对于每一个战斗型来说是 150 卡路里。画出新的支付表并且找出此时该种群中的所有 ESS。在这里进行的是什么类型的博弈？
   (c) 运用 13.6 节中的鹰鸽博弈概念，给出 (a) 和 (b) 中的 V 和 C 值，且证明你关于这些部分的答案与书中给出来的分析是相一致的。
3. 假设一次性囚徒困境博弈的支付情况如下：

		参与人 2	
		合作	背叛
参与人 1	合作	3，3	1，4
	背叛	4，1	2，2

在一个大型种群中，每个成员的行为由基因决定。每个参与人要么是背叛者（即在任何一次囚徒困境博弈中总是背叛），要么是以牙还牙的参与人（在多轮的囚徒困境博弈中，他在第一轮中合作；之后的每一轮中，他将选择其对手在上一轮中所选择的策略）。从这个种群中随机挑选出来的若干对参与人进行若干"盘"n次囚徒困境博弈（其中$n\geqslant 2$）。参与人每盘（进行n次）的支付为他在n次博弈中的支付之和。

设种群中背叛者的比例为p，则以牙还牙参与人的比例为$(1-p)$。每盘开始时，种群中的每个成员重复地进行若干盘囚徒困境博弈，在新的每一盘中都与随机挑选出来的新对手进行对局。一个以牙还牙参与人总是在其第一轮里以合作开始新的一盘。

(a) 在一盘博弈里，每个参与人与两种类型中的每一种对手相遇时，在一张2×2的表中表示出每种类型的参与人的支付。

(b) 计算背叛者的适应性（即参与人在整盘中与随机挑选出来的对手之间对局的期望支付）。

(c) 计算以牙还牙参与人的适应性。

(d) 利用（b）、（c）的答案证明，当$p>(n-2)/(n-1)$时，背叛型的适应性更强，当$p<(n-2)/(n-1)$时，以牙还牙型的适应性更强。

(e) 如果进化导致种群中有较强适应性的类型比例逐渐增加，那么对于这道习题中所描述的种群来说，什么是该过程可能的最终均衡结果？（即什么是可能的均衡，以及哪一些是进化稳定的？）用适应性图形来说明你的答案。

(f) 在什么样的意义上很多次的重复（n的数值很大）会促成合作的进化？

4. 在13.7.1节中，我们考虑了在三次重复的餐馆定价囚徒困境博弈中的ESS检验。

(a) 完整地解释（在表13—9的支付结构基础上）为什么一个完全由A类型组成的种群不可能被N类型或T类型变异所侵入？

(b) 解释为什么一个完全由N类型组成的种群会被A类型和T两种类型的变异所侵入？

(c) 最后解释为什么一个完全由T类型组成的种群不会被A类型的变异所侵入但会被N类型的变异所侵入？

5. 考虑一个种群，其中存在着两种表现型：天生的合作者（在审问下总是不会承认的）和天生的背叛者（准备好要承认的）。如果该种群的两个成员是随机地挑选出来的，他们在单轮博弈中的支付与第4章里的丈夫—妻子囚徒困境博弈中的支付相同，这里重新在下面列出来：

		行	
		承认	不承认
列	承认	10 年，10 年	1 年，25 年
	不承认	25 年，1 年	3 年，3 年

(a) 假设在种群中存在两种类型的可供选择的策略，就像在 13.2 节中的餐馆囚徒困境博弈中那样。这两个策略是 A（总是承认）和 T（以牙还牙，开始时不承认）。进一步假设每一对被挑选出来的参与人相继地进行两次囚徒困境博弈。画出这个两次重复囚徒困境博弈的支付表。

(b) 找出这个博弈的所有 ESS。

(c) 现在再增加第三个可能的策略 N，即绝不承认。画出这个具有三个可能策略的两次重复囚徒困境博弈的支付表，并且找出该博弈的这个新版本的所有 ESS。

6. 在习题 3 中的两次重复囚徒困境博弈中，假设存在着第四种可能的类型（类型 S），它也可以存在于种群里。该类型在第一轮里选择不承认，并且在与同样对手的相继两次对局中，每一次的第二轮里都选择承认。

(a) 画出这个博弈的 4×4 适应性表格。

(b) 如果该种群最初有这四种类型的每一种中的一些个体，证明类型 N 类型，T 类型，以及 S 类型会按照上述顺序相继消亡。

(c) 证明当某一策略（N，T，或 S）已经消亡时，正如在（b）中那样，倘若此种类型的一个很小比例的变异在未来重新出现，它们在那时不可能成功地侵入该种群。

(d) 该博弈的 ESS 是什么？为什么它不同于习题 5 的（c）中所发现的结果？

7. 考虑 13.7 节中重复三次的囚徒困境。在其中，有三种类型的餐馆老板参与定价博弈。就该博弈重做习题 4。

8. 回忆习题 2 中的为了价值为 200 卡路里的一种食物源进行战斗的一个种群。假定，正如在习题 2（b）中那样，每个战斗型一次战斗的成本是 150 卡路里。还假定在种群中存在着第三种表现型，它是一种混合体，所采取的是混合策略，有时战斗而有时又与人共享。

(a) 就像在理性决策博弈中找出混合策略一样，运用你的知识，为该博弈中使用混合策略的表现型设定一个合理的混合策略。

(b) 当采取混合策略的表现型采取了你在（a）里发现的混合策略时，画出该博弈的 3×3 的支付表。

(c) 确定这种混合表现型是不是该博弈的一种 ESS。（提示：你一定要检验一个混合表现型种群是否会被战斗型或共享型变异所侵入。）

9. 在本章的安全（约会地点）博弈中，支付代表参与人在不同的结果下所获得的物质的价值，例如，它可以是一次成功约会所给予的奖励。那么，种群中的其他个人可观察到这两种类型的期望支付（适应性），看看孰高孰低，并逐渐地模仿适应性强的策略。因此，种群中的两种类型的比例会发生改变。但

我们还可以从更加生物学的角度来解释这种结果，假设行参与人总为女性，列参与人总为男性，于是两类参与人因成功相遇而配对繁殖，并且他们的孩子是与父母相同类型的。因此，这些类型会作为相遇的一个结果而不断存续发展或者消亡。该博弈的这个新版本的正式数学描述使得其成为一种“双物种博弈”（虽然其在生物学上并非如此）。因此，于是种群中S类型女性的比例——称这个比例为 x——不需要等于种群中S类型男性的比例——称这个比例为 y。

试运用类似于性别战博弈中所应用的方法去研究 x 和 y 的动态变化，并找出稳定结果或该动态过程的结果。

10.（选做）在13.7.2节表13—10中的三种类型进化博弈中，设 $q_3=1-q_1-q_2$ 代表橙色颈前部攻击类型的比例。那么，每一种类型蜥蜴的种群比例的动态过程可以表达为

$$q_1 \text{ 增加当且仅当 } -q_2+q_3>0$$

和

$$q_2 \text{ 增加当且仅当 } q_1-q_3>0$$

我们在书中没有明确把它说出来，但是对于 q_3 的一个类似的法则就是

$$q_3 \text{ 增加当且仅当 } -q_1+q_2>0$$

(a) 更加明确地考察动态过程。设变量 x 随时间 t 的变化速度用导数 dx/dt 表示。则假定有

$$dq_1/dt=-q_2+q_3, \quad dq_2/dt=q_1-q_3$$

以及

$$dq_3/dt=-q_1+q_2$$

验证这些导数与前面有关种群动态过程的陈述是一致的。

(b) 假定 $X=(q_1)^2+(q_2)^2+(q_3)^2$ 运用微积分的知识证明 $dX/dt=0$，即证明 X 不随时间变化而变化。

(c) 从题干的定义中我们知道有 $q_1+q_2+q_3=1$。将这一事实与（b）中的结果相结合，证明随着时间的推移，在三维空间中，点 (q_1, q_2, q_3) 沿一个圆周运动。

(d) 在有色颈前部蜥蜴种群的进化动态过程的稳定性方面，由（c）的答案可以得出什么结论？

【注释】

[1] 表现型的动态变化由遗传型的潜在动态变化所驱动，但是，至少是在初步水平上，进化生物学将其分析集中在表现型水平上并且忽略进化的遗传方面。我们在进化博弈的研究上也采取类似的做法。在注释2引用的资料里可以找到遗传型水平上的一些理论。

[2] Robert Pool, “putting game Theory to the Test,” *Science*, vol. 267, 1995, pp. 1591 - 1593。这是近来出现的一篇适合普通读者阅读的好文章，其中有许多来自生

物学方面的例子。John Maynard Smith 在他的《进化遗传学》（*Evolutionary Genetics*, Oxford: Oxford University Press1989, chap. 7）和《进化与博弈论》（*Evolution and Game Theory*, Cambridge: Cambridge University Press, 1982）中讨论了生物学中的这种博弈；前者还介绍了许多进化的背景知识。我们向高级读者推荐 Peter Hammerstein and Reinhard Selten, "Game Theory and Evolutionary Biology," in *Handbook of Game Theory*, vol. 2, ed. R. J. Aumann and S. hart, Amsterdam: North Holland, 1994, pp. 929 - 993; and Jorgen Weibull, *Evolutionary Game Theory*, Cambridge, Mass: MIT Press, 1995。

[3] 事实上，进化论观点的应用不仅仅止于博弈论。下面这则笑话就给出了一种"引力的进化理论"来作为对牛顿或者爱因斯坦理论的一种替代。

问：为什么苹果会从树上掉到地上？

答：本来从树上掉下的苹果可以沿着任何方向飞出，但是只有那些遗传性地掉向地球的苹果才会再生出来。

[4] 在前面的章节中所讨论的理性行为里，我们会说背叛是一个严格占优策略。

[5] 该博弈只是两次重复博弈的一个例子。在基本博弈的其他支付下，两次重复不会有不明确的答案。第 4 章中的丈夫—妻子囚徒困境就是如此，见本章习题 5。

[6] 精确地说，种群里任何特定类型的比例都是有限的，且仅仅能够取诸如 1/1 000 000，2/1 000 000 这样的数值。但是，倘若种群足够大，并且我们在表 13—3 中用一条直线上的点来表示所有这样的数值，则这些点会非常紧密地聚集在一起，而且我们可将它们看成是一条连续的线。这就使得比例可取 0 到 1 之间的任何实数，从而我们就可考察特定行为类型在种群中的比例。根据同样的理由，某个成员如果入狱且被排除在种群之外，不会改变各种各样表现型在种群中的比例。

[7] 你现在应该为两次重复的情形画出一个类似的图形。你会看到对于所有小于 1 的 x 值，A 线是位于 T 线的上方，但两者在图的右边 $x=1$ 处相遇。

[8]《懦夫变异的入侵》(*The Invasion of the Mutants Wimps*) 将会是一部有趣的科幻喜剧电影。

[9] 也可以存在这样的进化稳定混合策略，其中种群中每一个成员采用一种混合策略。我们在 13.6.5 节中将进一步探讨这样一种想法。

[10] 在进化生物学里，这种类型的博弈被称为"不对称"博弈。对称博弈中一个参与人不能够简单地从观察另外一个参与人的外部特征来区分其类型；在不对称博弈里，参与人相互之间都能够识别。

[11] 而且这个发现支持且为混合策略均衡的性质提供了一种不同的理解，每一个参与人的混合策略使得其他的参与人在其纯策略之间是无差异的。现在我们可以将其看做，比如说，在一个两物种博弈的一个多态型进化均衡中，每种物种类型的比例使得另一物种所有幸存类型具有相同的适应性。

[12] 如果要了解有关单面斑点蜥蜴更多的信息，可参见 Kelly Zamudio and Barry Sinervo, "Polygyny, mate-Guarding, and Posthumous Fertilizations As Alternative Mating Strategies," *Proceedings of the National Academy of Sciences*, Vol. 97, no. 26, 2000, pp. 14427 - 14432。

[13] 第 7 章中的习题 4 考虑了 RPS 博弈的一种版本的理性均衡。你应该能够相对容易地验证这样的结论，即该博弈没有纯策略均衡。

[14] 要稍加注意以确保这三个比例的总和为 1，且在这个过程中，我们并不需要使用数学方法，而是用一种简单的方式表达我们的想法。本章习题 10 要求那些具有充分数学训练的读者给出更加严格的动态分析。

[15] 关于大肠杆菌的研究参见 Martin Nowak and Karl Sigmund, "Biodiversity: Bacterial Game Dynamics," *Nature*, Vol. 418, 2002, P. 138。如果这三种系在常规基础上被强制性地分散开，其中一种系会在大约几天的时间里占据统治地位并繁衍出第二种系，而第二种系将会迅速地杀死第三种系。

[16] 相反地，想要了解更多细节的读者可以读读 Maynard Smith 的《进化与博弈论》(*Evolution and the Theory of Game*)，特别是其中的第 14 页和第 15 页。John Maynard Smith 是进化博弈论的先驱。

[17] 倘若种群最初是多态型的且 m 是 J 类型的比例的话，则 m 不会是"很小的"。然而，m 的大小不再是一个关键性的问题，因为 $W(I)—W(J)$ 中的第二项现在被假定为零。

[18] 参见他在 *Cheating Monkeys and Citizen Bees: The Nature of Cooperation in Animals and Humans* (Cambridge: Harvard University Press, 1999) 中的精彩阐述。

[19] 在此处非常简单的讨论中，我们不可能对所有的议题和争论进行评论。一本非常流行的著作，也是本节里的许多例子的来源，是 Matt Ridley 的《美德的起源》(*The Origins of Virtue*, New York: Penguin, 1996)。我们还要指出的是，我们既没有在任何细节上考察遗传型与表现型之间的联系，也没有考察性在进化中的角色。Ridley 所写的另外一本书《红色皇后》(*The Red Queen*, New York: Penguin, 1995) 给出了对这个问题的一种非常精彩的处理。

[20] Dugatkin, *Cheating Monkeys*, p. 99.

[21] 根据强调基因水平上的选择的严格进化论，群体性的利他主义通常被认为是不可能的，但是在更为复杂的情况下，这个概念是可以得到解释的。更详细的讨论可参见 Dugatkin, *Cheating Monkeys*, pp. 141 - 145。

[22] Apsley Cherry-Garrard, *The Worst Journey in the World*, London: Constable, 1922; reprinted New York: Carroll and Graf, 1989, pp. 485 - 486.

[23] 关于利他性惩罚的证据，可以参见 Ernst Fehr and Simon Gachter, "Altruistic Punishment in humans," *Nature*, Vol. 415, 2002, pp. 137 - 140。

[24] Ridley, *The Origins of Virtue*, p. 83.

经济科学译库 · 策略博弈
经济科学译库 · 策略博弈 经济科学译库 · 策略博弈 经济科学译库 · 策略博弈

第四部分

在特定策略情形下的应用





第 14 章　边缘政策：古巴导弹危机

在第 1 章中，我们解释过我们的基本方法既不是纯粹的理论也不是纯粹的案例研究，而是这样的一种组合，其中理论的思想是通过利用特定的案例或者例子的特征而发展出来的。因此，我们将每一个案例中那些对于被发展起来的概念是偶然和次要的方面忽略掉。然而，在学习了理论思想之后，你就有了一种十分丰满的分析模型。在这种模型中，特定案例的事实性细节就更加紧密地与博弈理论的分析整合在一起了，从而对于发生了什么和为什么发生达成一种充分的理解。这样的基于理论的案例研究（theory-based case studies）已经出现在许多领域如商业、政治科学以及经济史中。[1]

在这里，我们提供一个来自政治和军事历史中的例子——发生于 1962 年古巴导弹危机中的核冒险政策。我们的选择是由下列因素促成的，即大量可以获得的事实信息，以及来自博弈论的一种重要概念的可应用性。你可以认为核战争的风险随着苏联的解体而消失了，并且因此我们的案例是一种历史猎奇。但是，核军备竞赛继续出现在世界上的许多地方，并且像印度与巴基斯坦或者伊朗与以色列这样的竞争对手都可以从古巴导弹危机中吸取经验。对于你们中的大多数人来说，更加重要的是，在许多日常性的场合里——例如，罢工和夫妻吵架——都必须运用边缘政策。尽管在这些博弈中的赌注比起超级大国之间的核对抗来说要低得多，但同样的策略性原理是适用的。

在第 10 章中，我们曾经将边缘政策概念作为一种策略性行动引入；这里只需要提示一下即可。威胁是一种反应规则，并且这种威胁行动给做出威胁的参与

人和威胁试图要去影响其行动的参与人两者都会带来成本。但是，如果威胁产生效力，则该行动实际上是不会被实施的。因此，对于威胁性行动的成本来说，是不存在明显的上限的。但出现错误的风险——也就是说，威胁可能不能产生效力或者威胁性行动可能由于偶然因素而实际发生——迫使策略家去运用能实现其目的的最小的威胁。如果一个较小的威胁不是本来就有的，那么可以通过降低威胁实施的可能性将一个大的威胁变小一些。你可以率先行动创造一种可能（当然不是确定性），当对手违背了你的意愿时，对双方都有伤害的结果就会产生。如果这种需要实际出现了，当你有充分的选择自由时你也不会采取那个坏的行动。所以你要事先为事态在一定程度上失去控制做好安排。边缘政策就是要创造和开发这样一种概率威胁，方式是有意地失去对事态的控制。

“边缘政策”一词的使用经常与核武器联系在一起。1962 年的古巴导弹危机使得世界距离核战争仅一步之遥。该危机经常被视为边缘政策的经典案例。我们也将对该危机进行案例研究去详细解释“边缘政策”这一概念。在这个过程之中，我们将发现对这场危机的许多流行的理解及分析实在是过于简单化了。较深入的分析揭示出边缘政策是一种巧妙地安排的和危险的策略。它也表明行动诸如罢工和关系破裂这样一些商业和个人之间的互动中的冲突都是边缘政策的运用出现错误的例子。因此，对于这种策略及其局限和风险的一种清晰的理解对于所有的博弈参与人来说都是非常重要的。当然，这对于每一个人来说也是相当重要的。

14.1 对事件的一个简要的叙述

我们从对该危机来龙去脉的简要描述开始。我们的材料来自几本书，包括一些利用由于苏联的解体而解密的文件和报告完成的著作。[2]我们并不期望能对细节做出公正的评价，而只是追随事件发展的脉络。肯尼迪总统在危机期间曾经说过“这是我挣得工资的一周”(This is the week when I earn my salary)，但当时让事件处于平衡状态要比总统的薪水重要得多。我们建议你们读读生动描述这一危机来龙去脉的书籍，并与“生逢其时”的亲戚们聊聊这一事件，以获取他们的“第一手”的记忆。[3]

1962 年夏末秋初，苏联（USSR）开始在古巴部署中程弹道导弹（MRBM 和 IRBM）。MRBM 有 1 100 英里的射程，可打到华盛顿特区；IRBM 的射程为 2 200 英里，能打击全美大多数主要城市和军事设施。部署导弹的阵地由最新式的苏式萨姆（SAM）地对空 SA-2 型导弹护卫，它能够击落美国的 U-2 高空侦察机。同时还部署了被称为 IL-28 的轰炸机和一种战术核武器。苏联人称这种战术核武器为 Luna，而美国人称其为青蛙（FROG，可地面移动的火箭）。古巴人可用这种武器抵挡入侵的军队。

这是苏联第一次试图把其导弹和核武器部署在领土之外的地方。倘若它成功了，将大大提高它对美国的攻击能力。现在已经清楚的是当时苏联境内可直接打到美国本土的已部署的洲际弹道导弹（ICBM）总共不超过 20 枚，甚至大概只有

2 到 3 枚（*War*，464，509－510）。他们最初在古巴部署了大约 40 枚 MRBM 和 IRBM，这是一个巨大数量的增加。但是，美国仍然在两个超级大国之间的核平衡上保持着巨大的优势。同时，由于苏联已经建造了自己的潜艇舰队，在美国本土附近部署陆基导弹的相对重要性已下降。但是，对于苏联来说，在古巴部署导弹的意义远远超出了直接的军事意义。成功地在如此接近美国本土的地方部署导弹对于苏联在全球范围内的声誉来说是一个巨大的提升，特别是在亚洲和非洲，两个超级大国在那里正在进行激烈的政治及军事影响的竞争。最后，苏联打算将古巴作为其“东方阵营”的一个桥头堡。通过在古巴部署导弹，以打消美国人侵古巴的念头，消除古巴所受到的来自美国的巨大威胁，同时，还抵消掉中国在古巴的影响，这在当时的苏联领导人尼基塔·赫鲁晓夫（Nikita Khrushchev）的盘算中具有十分重要的分量（对苏联动机的分析，参见 *Gamble*，182－183）。

在 1962 年夏末秋初期间，美国对古巴及海上行船航线进行监视的系统发现了一些可疑的活动。当美国外交官向苏联方面寻求这些疑问的答案时，苏联立即对在古巴部署导弹的任何意图进行了否认。随后，面对难以掩饰的证据，他们才说其意图是防守性的，是为了防止美国对古巴的侵略。尽管进攻性武器也可用于防卫性威慑，但这种解释显然不能令美方信服。

在 10 月 14 日和 15 日，即星期天和星期一，一架美国 U-2“间谍飞机”拍摄了古巴西部地区的照片。当把胶卷冲洗出来并且加以识别后，它们无可辩驳地显示出正进行着 MRBM 发射基地的建造（随后在 10 月 17 日又找到了 IRBM 的证据）。这些照片在随后（10 月 16 日）被提交给肯尼迪总统。他立即组建了一个特别顾问小组，这个小组后来被称为国家安全委员会执行委员会（Executive Committee of the National Security Council，简称 ExComn）。他们开始讨论对策。在第一次会议上（10 月 16 日上午），肯尼迪总统决定在他完全准备好应对方案之前，对这件事实行完全地保密和消息封锁。他这样做的主要原因是如果苏联知道美国知道了古巴的事情，会赶在美国做出反应之前加快在古巴的导弹安装及部署。除此之外，这样做也可以避免在美国境内造成人们在政府没有明确的反应下出现大恐慌。

国家安全委员会执行委员会中积极参与对策建议的成员包括国防部长罗伯特·麦克纳马拉（Robert McNamara）、国家安全顾问乔治·邦迪（George Bundy）、参谋长联席会议主席（the Chairman of the Joint Chiefs of Staff）麦克斯威尔·泰勒将军（General Maxwell Taylor）、国务卿迪安·腊斯克（Dean Rusk）以及副国务卿乔治·波尔（George Ball）、首席检察官罗伯特·肯尼迪（Robert Kennedy，他也是肯尼迪总统的兄弟）、财政部长道格拉斯·迪隆（Douglas Dillon，他是内阁中唯一的共和党人），以及莱威林·汤普森（Llewellyn Thompson，他是不久前才从驻莫斯科大使任上返回美国的）。在随后的两个星期里，他们与其他的一些重要官员一起商讨应对措施。这些官员包括美国驻联合国大使阿德莱·史蒂文森（Adlai Stevenson）、前国务卿和美国外交政策的高级发言人迪恩·艾奇逊（Dean Acheson），以及美国空军司令柯蒂斯·李梅将军（General Curtis Lemay）。

在这一周剩下的时间里（10 月 16 日到 21 日），国家安全委员会执行委员会

进行了多次的集中性讨论。为了保密，总统继续了他原有的正常计划日程，包括为即将参加在1962年11月举行的国会中期选举中的民主党候选人做巡回演说。当然，他仍然暗地里保持着与国家安全委员会执行委员会的经常性接触。他躲避着报刊关于古巴问题的采访，并且劝说一两个可信任的媒体老板或编辑保持一贯的经营状态。国家安全委员会执行委员会自身在试图在华盛顿保守秘密的做法上有时表现得接近于滑稽。比如有一次他们中的12个人像沙丁鱼一般挤在一辆轿车里，因为由几辆政府小车组成的长长的车队在护卫车的簇拥下从白宫向其他政府部门行驶的景象将会在媒体中引起猜测。

国家安全委员会执行委员会中的不同成员在对时局的估计上存在着广泛的分歧，各自支持采取不同的行动。参谋长联席会议主席认为苏联在古巴部署导弹大大打破了军事力量的平衡；国防部长麦克纳马拉认为它"根本"没有影响到军事平衡，并且仍然认为这只是一个纯政治的问题（*Tapes*，89）。总统肯尼迪指出，如果美国对苏联第一次导弹部署不做出反应，苏联将会在随后部署更多的导弹；并且苏联将会利用在美国的后院部署导弹所造成的威胁迫使美国将在西柏林驻守的美军、英军及法军撤出。肯尼迪也意识到它是美苏之间地缘政治斗争的一个组成部分（*Tapes*，92）。

现在看来他的估计是非常正确的。苏联计划将它在古巴的存在进一步扩充为一个主要的军事基地（*Tapes*，677）。他们打算在11月中旬之前就完成导弹的部署。赫鲁晓夫已经计划在11月的下旬与卡斯特罗签订一个协议，然后就去纽约的联合国总部发表演讲，接着打算发出一个有关柏林问题的最后通牒（*Tapes*，679；*Gamble*，182），利用在古巴部署的导弹作为达到这一目的的要挟。赫鲁晓夫认为肯尼迪将不得不将已部署的导弹作为一个既成事实来接受。现在已清楚的是，赫鲁晓夫当年是私下独自一个人制定这些计划的。他的一些高级幕僚私下认为他们已过于冒险了，但是苏联的最高政府决策团体苏联最高苏维埃主席团却支持他的决定，尽管它本身只是一个橡皮图章（*Gamble*，180）。卡斯特罗开始并不愿意接受导弹部署，因为他害怕由此引起美国的入侵（*Tapes*，676－678），但后来他还是接受了，并且未来前景给予他自信心，让他在有关美国的演说中更加趾高气扬（*Gamble*，186－187，229－230）。

包括10月18日即星期四上午的那次会议在内，在之前的国家安全委员会执行委员会召开的所有会议上，每个人都认为美国的反应将会是单纯的军事行动。在此期间，他们认真研究过的全部选择是：(1）专门针对导弹阵地及邻近的萨姆导弹阵地（有可能存在）进行一次直接空袭；(2）对包括苏联和古巴停在机场上的所有飞机的一次更广泛的空袭；及（3）对古巴的全面入侵。如果需要的话，当射程更远的IRBM存在的证据取得后，态度还可更强硬。事实上，在星期四的会议上，肯尼迪讨论了在周末开始进行空袭的时间表（*Tapes*，148）。

麦克纳马拉在10月16日礼拜二会议结束前首次提到了军事封锁的计划，并且在正式的会议结束后在小范围内提出了这种思路（在实际行动将要发生的时候以一种神秘的方式）（*Tapes*，86，113）。波尔指出在没有预先发出警告的情况下进行空袭无异于另一次"珍珠港"事件，并且美国人是不愿干此勾当的（*Tapes*，115）；他得到了罗伯特·肯尼迪的强力支持（*Tapes*，149）。国家安全委员会执

行委员会中的非军人成员们在他们发觉参谋长联席会议成员们需要的是一次大规模空中打击时，就进一步地倾向于军事封锁的选择；认为军方将目标仅限于对导弹阵地的有限打击是危险且无效的，以致“他们宁愿选择不采取任何军事行动也不愿采取有限打击的方式”（Tapes，97）。

在 10 月 18 日与 20 日礼拜六之间，国家安全委员会执行委员会逐渐取得了一致意见，即先进行封锁，同时发出一个具有较短截止期限（48 小时到 72 小时）的最后通牒，在截止期限结束时，如果有必要，再采取军事行动。根据国际法，实行封锁之前需要先宣战，但这一问题通过将行动命名为对古巴的“海上隔离”（naval quarantine）而巧妙地得到了解决（*Tapes*，190－196）。

在这些讨论（从 10 月 16 日到 21 日）中，一些人总是持有相同的立场——例如，军官们坚持进行一次大规模的空袭，但是其他人却随时都在改变他们自己的观点，有些时候甚至极富戏剧性。邦迪最初赞成“什么都不干”（*Tapes*，172），但后来却转向于一种先发制人的空中打击方案（*Tapes*，189）。肯尼迪总统本人的立场也从空袭转向军事封锁。他要求美国的反应是坚决的。虽然他的理由无疑主要是军事及地缘政治的，但作为一位成熟的国内政治家，他充分地意识到一种不够强硬的反应在即将举行的国会选举中将不利于民主党。另一方面，启动一次可能导致核战争的行动所应负的责任使他深感压力沉重。中央情报局的评估认为一些导弹已经部署完毕。这就增长了任何空袭或入侵引发苏联人发射这些导弹并造成美国平民大量伤亡的风险（*Gamble*，235）。肯尼迪对此评估印象深刻。在危机持续到第二个星期时（10 月 22 日到 28 日），他的决定看起来是坚持选择由国家安全委员会执行委员会讨论过的最保守的方案。

第一周的讨论结束之后，选择在军事封锁与空袭之间进行。持两种立场的报告都递交上来，并且在 10 月 20 日的一次非正式投票中，军事封锁建议以 11 比 6 取胜（*War*，516）。肯尼迪做出了抉择并且于 10 月 22 日在电视上发表演说，向全美公众宣布军事封锁行动开始了。他要求苏联向古巴运送导弹的船只立即停止航行，同时迅速撤除已经部署在古巴的导弹。

肯尼迪的演说将戏剧性的事件引发的紧张的局势带入公共领域。联合国展开了几次充满戏剧性但却无效的商讨和争论。其他的世界领导人和职业性的国际事务协调人也纷纷从中斡旋，并提出了若干建议。

从 10 月 23 日到 25 日，苏联人咆哮着试图恫吓和否认，赫鲁晓夫称军事封锁为“盗贼行为，是一种愚蠢的国际帝国主义行为”，并且声称他的船只将不会理睬它。苏联人在联合国或其他地方到处宣称他们的意图只是防御，并且发表了蔑视性的声明。但在私下里，他们却在寻找结束危机的渠道。赫鲁晓夫曾向肯尼迪直接发出过一些信息，此外苏联人还采用过一些间接和较低层次的沟通渠道。事实上，早在星期一，即 10 月 22 日，在肯尼迪发表电视演说之前，最高苏维埃主席团已决定不要将这次危机引向战争。直到礼拜四，即 10 月 25 日，他们才决定从古巴撤回导弹，但要求作为一种交换，美国要做出不入侵古巴的承诺。但是，他们也同意“再看一看”（Look around），以便寻求到更好的办法（*Gamble*，241，259）。当时美国并不知道苏联人的这种思想准备。

在公开场合以及私下的通信里，苏联人提及了用一种双方都做出让步的办法

来解决危机的可能性，即美国从土耳其撤出导弹，苏联从古巴撤出导弹。在国家安全委员会执行委员会里也已对这种可能性做了讨论。美国在土耳其的导弹是已经准备废除的设施，因为美国正打算将它们撤走并用游弋在地中海底下的北极星潜艇去取代它们。但是，可能土耳其人很难被说服接受这一改变，因为他们将美国导弹在其领土的存在作为对本国的一种保护性象征（土耳其人的想法是正确的，因为固定在土耳其领土上的导弹是对其防卫的一种强大的承诺性象征，比起在其海岸线之外游弋的潜艇来说，其对土国防卫的承诺力量更大，因为潜水艇在很短的时间就可能逃之夭夭，参见 *Tapes*，568）。

在 10 月 24 日即星期三那一天，军事封锁开始发挥作用了。尽管苏联人在公共场合仍然表现得十分强硬，但在处理这一事件时开始变得小心翼翼了。看来，他们对美国人能够在导弹部署安装整个完成之前就发现在古巴有导弹感到惊讶；苏联在古巴的工作人员已经发现了 U-2 飞机在飞越古巴领空，但他们并没有向莫斯科报告（*Tapes*，681）。最高苏维埃主席团下令所有载有最敏感物质（实际上就是 IRBM 导弹）的船只停下来并调转方向返回。但他们又命令驻古巴的苏军司令伊萨·普利耶夫将军（General Issa Pliyev）下令他的部队处于战斗准备状态，并且要求他使用除核武器外的所有可能的手段迎击来犯之敌（*Tapes*，682）。事实上，最高苏维埃主席团曾两次打算但又在发出命令之前取消授予他使用战术核武器迎击可能来犯的美军的权力（*Gamble*，242－243，272，276）。当然，美方仅仅看见几艘继续驶向军事封锁区的苏联船只，这些船只实际上运载的是油料和其他的非军用货物。美国海军在执行其军事封锁任务时开始转向缓和。有一艘油船未经登船检查就让它通过了，而另一艘运有工业货物的不定期汽船经粗略的登船检查之后就让其继续前行。但是，紧张的状态仍然在继续升级，并且双方的行为都比双方最高层政治家所希望的来得过火一些。

在 10 月 26 日礼拜五的早晨，赫鲁晓夫私下给肯尼迪写了一封和解信，其中提出只要美国承诺不入侵古巴，他就将导弹撤出。但是，第二天他的态度又强硬起来。看来，他被所看到的两个现象壮了胆子。第一是美国海军在执行军事封锁时没有过于挑衅的行为。他们让一些显然是民用的货船通过封锁区，只是登上一艘叫 Marucla 的船进行了检查，在进行大而化之的登船检查后就让它通过了。第二是在美国国内的报纸上出现了一些带有点鸽派味道的言论。其中最为引人注目的是由著名的专栏作家沃特·利普曼（Walter Lippman）所写的一篇文章。他在这篇文章中建议美国以在土耳其撤出其导弹来换取苏联撤出其在古巴部署的导弹（*Gamble*，275）。赫鲁晓夫于是在 10 月 26 日即星期六[1]给肯尼迪写了第二封信，提出了这一主张，并且这一次他把信的内容公开了。这一封新的信大概就是最高苏维埃主席团所主张的最佳处理方式中的“再看一看”（Looking around）策略的一个组成部分。国家安全委员会执行委员会得出的结论是第一封信表达了赫鲁晓夫本人的想法，而第二封则是在最高苏维埃主席团中强硬派的压力之下写的——或者甚至可能是赫鲁晓夫此时已不再拥有控制权的证据（*Tapes*，498，512－513）。

事实上，后来发现，两封信都是经最高苏维埃主席团讨论和批准之后发出的

① 原书如此，疑有误。——译者注

(*Gamble*，263，275)。

国家安全委员会执行委员会的成员们继续商讨对策，并且大家的意见也开始变得强硬起来。一种理由是大家逐步感觉到仅仅依靠封锁是不能解决问题的。肯尼迪的电视讲话没有给出一个明确的期限，并且，正如我们已知道的那样，一个没有规定明确期限的强迫性威胁是易于因对手故意拖延而被削弱的。肯尼迪早在10月22日即星期一就对此很清楚了。在肯尼迪发表演说之前的一次国家安全委员会执行委员会晨会上，他评论道："我不认为让他们一直待在那儿，我们的境况就会好起来"(*Tapes*，216)。但是，一个刚性的、很短时间的期限似乎又太缺乏灵活性了。到了礼拜四，国家安全委员会执行委员会中的其他人也开始意识到这一问题。例如，邦迪说道："等待是最危险的了"(*Tapes*，423)。苏联人的立场变得强硬起来，正如在私下发出的安抚性"周五信件"之后的公开性的"周六信件"所表明的那样，这是另外需要关注的问题。最为岌岌可危的是，在星期五，美国的情报部门已发现在古巴部署有战术核武器(FROG)(*Tapes*，475)。这一发现表明苏联在那儿的存在比早先所想象的还要严重得多，但这也使得入侵变得对美军十分不利和危险。还是在星期六，一架美国U-2飞机在古巴上空被击落(现在已知道，这是地区司令官干的，他把上面的命令理解得比莫斯科的意愿宽泛得多了)(*War*，573；*Tapes*，682)。另外，古巴防空系统向低空飞行的美国侦察机开了火。国家安全委员会执行委员会在那个星期六的郁闷心情可用迪隆的一句话来概括："我们这一天什么也没有得到"(We haven't got but one more day)(*Tapes*，534)。

在星期六，一项导致紧张局势逐步升级的计划开始被付诸行动。在随后的礼拜一，或最迟是礼拜二，将实施空中打击，并且对空军预备队下达了动员令(*Tapes*，612-613)。势态发展的必然性终结点将是入侵(*Tapes*，537-538)。一封肯尼迪总统私下致赫鲁晓夫的强硬信件已经起草，并且由罗伯特·肯尼迪直接送交到苏联驻华盛顿大使安东尼·多勃雷宁(Anatoly Dobrynin)手中。在这封信中，肯尼迪提出以下建议：

(1) 苏联将部署在古巴的导弹和IL-28轰炸机撤出，并且要进行核查(且不许有新的船只驶入古巴)。

(2) 美国承诺不入侵古巴。

(3) 几个月后美国从土耳其撤出自己的导弹，但若苏联在公开场合提到它或者将此行动与古巴危机相联系的话，美国将放弃这一行动。他要求苏联在12小时到24小时内给予答复，否则将会有灾难性的后果(there would be drastic consequences)(*Tapes*，605-607)。

在星期天早上，即10月28日上午，正当美国的许多教堂内挤满了祷告和祈求和平的人们时，苏联的无线电广播电台播送了赫鲁晓夫回复给肯尼迪的一封信的内容，其中他声明将立即停止导弹阵地的建造工作，并且将拆除已安装好了的导弹并且装船运回苏联。肯尼迪立即对此加以回应，表示对此决定大加欢迎，这一回应由美国之音向莫斯科播送。现在已经清楚的是，赫鲁晓夫做出让步的决定实际上在他收到多勃雷宁转交的肯尼迪的信件之前就已做出了，但是这封信使得赫鲁晓夫更加毫不犹豫地做出了这一决定(*Tapes*，689)。

但这并未完全终止这场危机。美方的参谋长联席会议怀疑苏联不会如其所说的那样撤出古巴，并坚持空袭（*Tapes*，635）。事实上，在后来的几天里，古巴导弹阵地上的建造活动还持续了一阵子，而联合国的检查也是成问题的。苏联也试图将美方将撤走部署在土耳其的导弹这一计划半公开化，他们还打算让 IL-28 轰炸机继续留在古巴。直到 11 月 20 日，危机才最终了结，导弹才开始被陆续撤出（*Tapes*，663－665；*Gamble*，298－310）。

14.2 一种简单的博弈论解释

初看起来，这场危机的博弈论解释是十分简单的。美国要求苏联从古巴撤出导弹；于是，美国人的目标是达到一种强迫性（Compellence）。为此，美国人有效地使用一种威胁：如果苏联人不顺从的话，将最终引发两个超级大国之间的一场核大战。军事封锁是这一必然过程的开始的第一步，并且是证明美国人威胁的置信度的实际行动。换句话说，肯尼迪将赫鲁晓夫推向战争灾难的边缘。这就充分地震慑了赫鲁晓夫并迫使其就范。当然，核毁灭的可怕后果同样也会令肯尼迪感到恐怖，但这正是威胁的本质。必需满足的一点是威胁一旦实施，将使对方付出惨重代价，从而迫使他们按己方意图行事；这样，我们实际上并不必实施这种极糟糕的行动。

我们通过画出一个博弈树，就可以将上述论断正式地表达出来，见图 14—1。

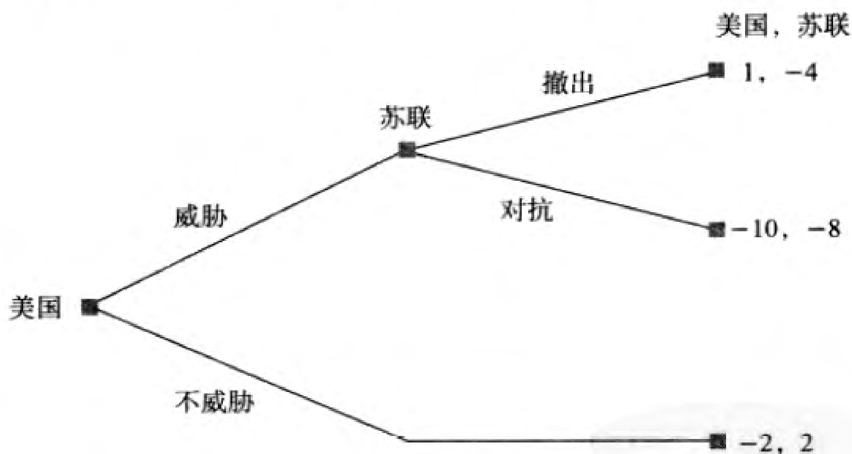


图 14—1　古巴导弹危机的简单威胁模型

苏联已经部署好了导弹，现在由美国首先行动。美国有两个选择，一是什么都不做（不威胁），二是发出威胁。倘若美国选择什么都不做，苏联将取得一个很大的军事及外交上的成就；故我们将美国此时的支付标记为−2，将苏联的支付标记为 2。如果美国发出威胁，就轮到苏联进行行动选择了。苏联只有两个选择——撤出或者对抗。对于苏联来说，选择撤出是令人丢尽脸面的（一个很大的负数），并且对于美国而言，其军事优势再一次得到证实（一个小的正数）；故我们将美

国的支付记为 1，将苏联的支付记为 −4。如果苏联对抗美国的威胁，将会爆发一场核战争。这对于双方来说都是十分可怕的，但对于美国人来说是更加恐怖的事，故将美国的支付记为 −10 而将苏联的支付记为 −8。这样的计算只是粗略的估算，但结论并不依赖于我们所选择的数字。如果你不同意我们选择的数字，你可以用任何你认为更具代表性的数字代替它们；只要结果的相对顺序不变，你会获得同样的子博弈完美均衡。

现在，我们可以很容易地找出子博弈完美均衡。如果面对美国的威胁，苏联从撤出获得的支付为 −4，从对抗中获得的支付为 −8，则他们更加偏好于撤出。给定这种预期，美国盘算着若发出威胁将获得支付 1，否则获得支付 −2；因此，对于美国来说，最优的选择是发出威胁。美国最终得支付 1，苏联得支付 −4。

然而，稍微深入一点地思考就会发现这一解释并不令人满意。人们会问，如果苏联人预见到在随后的博弈中他们将成为输家，那么他们为什么还会在古巴部署导弹呢？更为重要的是，关于当时情形的一些事实以及博弈展开过程中的一些事件与该简单威胁的图示并不吻合。

但是，在对这种分析的不足做出说明并提出一种更好的解释之前，我们暂且离开主题去看看危机期间发生的一个有趣的情节，它使得我们更加了解一个成功的强迫性威胁到底需要什么样的附加条件。正如在第 10 章中所指出的那样，一个强迫性威胁必须要有一个最后期限，否则对手能通过拖延战术来使其无效。在 10 月 23 日即星期二的联合国秘书处的危机讨论会上，美国大使阿德莱·史蒂文森与苏联驻联合国代表瓦列里安·佐林（Valerian Zorin）发生了冲突。史蒂文森要求佐林直截了当地回答苏联是已经在古巴部署了还是正在部署核导弹——"'是'或'不是'——不要等翻译——'是'还是'不是'？回答我！"他坚持地要求道。佐林反驳道："我并不是站在美国的法庭上……你会随着事情的发展自然获得答案的。"史蒂文森又反击道："我打算等你的答复，一直等到世界末日"(I am prepared to wait for my answer until hell freezes over)。这一场争吵极富戏剧色彩。肯尼迪看着电视上的会议实况，评论道："真棒！我还从来不知道阿德莱还真有他的！"（*Profile*，406）。但是，这是一个可怕的策略。当苏联人正忙于建造其导弹阵地时，没有什么比让美国人"等着答复"更合苏联人的胃口了。对于强迫性威胁来说，"一直等到世界末日"并不是一个合适的期限。

14.3 考虑更多复杂性后的博弈

让我们回过头来构造一个更加令人满意的博弈论模型。正如我们在前面所指出的那样，一个威胁在其大小上仅有下限——即它是足够大的足以震慑对手——的思想仅仅在威胁者能够绝对肯定地保证所有事情都按照计划进行时才是正确的。但是几乎所有的博弈都存在某些不确定的因素。你不可能确切地知道你的对手的价值观，也不可能完全肯定参与人打算采用的行动都会被准确无误地贯彻。因此，一个威胁会有两方面的风险。你的对手可能会对抗，使你不得不在蒙受很

大损失的情况下执行威胁行动；或者你的对手会选择顺从，但威胁行动也可能由于失误而发生。当这种风险出现时，威胁行动给自己带来的成本将成为一个需要首先考虑的问题。

古巴导弹危机期间就充满着大量的这类风险。任何一方都不能准确知悉另一方的支付，即另一方将核战争和在国际上丢脸的相关成本看得有多么大。还有，选择“军事封锁”和“空袭”比简单的口头承诺要复杂得多，并且在华盛顿或莫斯科发出的指令与其在大西洋上或古巴的具体执行之间存在许多信息失真和随机效应。格雷汉姆·阿里森（Graham Allison）写的一本十分优秀的著作《决策的本质》（*Essence of Decision*）曾对所有的这些复杂性和不确定性加以了概括。这使得他得出如下的结论：“古巴导弹危机不可能用博弈论来解释。”他考虑了两种可能的解释：一种是基于这样的一个事实，即官僚政治有着自己的行事规则和程式；另一种则是基于美国的内部政治和苏联的政府控制与军事机器的特点。他由此认为：用政治来解释才是最合适的。

我们大体上同意他的观点，但对古巴导弹危机有着不同理解。并不是博弈论不能用来理解和解释古巴导弹危机，而是这次危机并不是一个二人博弈——美国与苏联或肯尼迪与赫鲁晓夫之间。双方本身其实都是一个由大量参与人组成的复杂联盟，这些成员都有着不同的目标、信息、行动和通讯方式。双方内部的成员都卷入了其他的博弈，并且某些成员也直接地与另一方的对手们交手。换句话说，这次危机可被视为一个复杂的多人博弈，参加博弈的人来自两个不同阵营。可将肯尼迪和赫鲁晓夫视为这个博弈中的最高级参与人，他们都受到一种制约，即都不得不在自己阵营中那些持不同观点和拥有不同信息的人中进行协调，并且对己方的那些人的行动都没有完全的控制力。我们将表明，这个更加精致的博弈论视角不仅对于理解危机是一种很好方式，而且在理解如何操作实际的边缘政策（冒险主义）博弈来说也是一种基本的方法。

我们从阿里森所强调的某些证据和来自其他作者的线索开始。

首先，每一方都有几种不同的建议。在美方，正如之前已经提到的，国家安全委员会执行委员会中的分歧很大。另外，肯尼迪发现有必要与诸如前总统艾森豪威尔及国会领导人等人商量。他们中的一些人有非常不同的看法，例如，参议员福尔布莱特（Fulbright）在一次私下聚会中说军事封锁“在我看来是最糟糕的选择了”（*Tapes*，271）。媒体和政治对手们也不会给予总统持久的无保留的支持。倘若幕僚们和公众的立场变成鹰派味十足了，肯尼迪就不再能继续他的稳健政策了。

在这两周之内，成员们也开始改变他们的立场。例如，麦克纳马拉在开始时是十分鸽派的，提出在古巴的导弹对于苏联威胁来说并无显著的增强（*Tapes*，89），并且倾向于军事封锁和谈判（*Tapes*，191），但后来却十分地鹰派，指出赫鲁晓夫在周五即10月26日发出的和解信是“充满漏洞的”（*Tapes*，495，585），并且提出要入侵古巴（*Tapes*，537）。最为重要的是，美方的军事将领们总是倡议来一场大规模的攻击性空袭。即使在危机结束并且每个人都认为美国已在冷战中赢得了重要的一个回合之后，空军将领李梅仍然不满意，并要求展开行动。“我们失去了机会！我们今天应该已占领了那里并将他们干掉了”，他说道（*Es-*

sence，206；*Profile*，425）。

尽管赫鲁晓夫是苏联的独裁者，他仍然无法完全控制住局势。苏联一方内部的意见分歧没有被完好地记录下来，但值得一提的是，后来人们的回忆指出赫鲁晓夫几乎是单独决定在古巴部署导弹的，并且当他告知其伙伴们时，他们认为这真是一场欠考虑的大赌博（*Tapes*，674；*Gamble*，180）。即使最高苏维埃主席团仅仅是一个橡皮图章，赫鲁晓夫跨越它进行决策在某种程度上也是受到一定的制约的。的确，两年之后，灾难性的古巴冒险就成为最高苏维埃主席团解除他职务的主要指控之一（*Gamble*，353－355）。有人说赫鲁晓夫本来打算对抗军事封锁，并且只是由于第一副主席阿纳斯塔斯·米高扬（Anastas Mikoyan）的坚持才使他做出了一种谨慎的反应（*War*，521）。最后，在10月27日即周六那一天，卡斯特罗命令他的防空部队向所有飞越古巴领空的美国飞机开火，并且还拒绝了苏联大使要求他取消这一命令的请求（*War*，544）。

在美方，不同团体有着非常不同的信息和对时局的非常不同的理解。并且，这会经常地引发一些行动。这些行动与领导人的意图是不一致的，甚至有时直接违反上面的命令。以摧毁导弹为目的的“空中打击”的概念就是一个很好的例子。国家安全委员会执行委员会中的非军人成员认为它只是一个非常狭隘的目标，并且不会造成古巴和苏联人的大量伤亡，但空军却要求展开更大规模的攻击。幸运的是，这些分歧出现得较早，从而国家安全委员会执行委员会决定放弃空袭且总统驳回了空军的提议（*Essence*，123，209）。至于军事封锁，美国海军为此已制定了行动。政治领导人要求一种不同的且较为柔性的过程：包围古巴，让苏联有更多的时间重新考虑，让明显是载有非军用物资的船只安全通过，并且对敢于对抗的船只只是使其失去战斗力，并不将其击沉。然而，除了麦克纳马拉的指引外，海军主要是按照其标准程序行事的（*Essence*，130－132）。美国空军甚至制造了很大的险情。一架U-2飞机盲目地飞行并“偶然地”进入了苏联领空，几乎导致严重后果。李梅将军，在没有总统的允许或者授权的情况下，下令美国战略空军的核轰炸机飞过其“返航点”并进入苏联领空中的一定距离去确定它们会在什么位置被苏联雷达探测到。幸运的是，苏联人的反应是冷静的，赫鲁晓夫仅仅是向肯尼迪提出了抗议。[4]

在苏方，在信息和通讯方面也存在类似的不足，并且指令和控制链条也很脆弱。例如，导弹的建造需通过标准的官僚程序来完成。苏联总是在他们自己的国土里那些不易于受到空袭的地方建造ICBM阵地，他们在古巴也是这样干的，而在那里他们更易受到袭击。在危机达到白热化的时候，当苏联萨姆-2导弹部队在10月26日即周五看见头上飞过的美国U-2飞机时，普利耶夫当时暂时不在办公室，他的副手下令将U-2飞机打下来；这一意外事件引发了远高于莫斯科所期望的风险（*Gamble*，277－288）。在其他的很多时候——例如，当美国海军试图截住货船Marucla号并打算登船时——一些人采取了一些行动几乎导致了一起骇人听闻的“事故”。甚至更具戏剧性的是，10月27日，一个苏联的潜艇水兵在接近于隔离线时曾经向水面部队发出警告，声称将考虑发射其所携带的一枚载有核弹头的鱼雷（美国海军不知道它携带有这样的武器）。发射授权规则规定发射鱼雷需要三名官员达成一致意见，而其中只有两名官员同意，第三名官员阻止了

一场核大战。[5]

所有这些因素使得双方最高领导的任何决策的结果都会出现某种不可预测性。这就导致了“威胁走向错误”（threat going wrong）的巨大风险。事实上，肯尼迪认为军事封锁会引发战争的可能性位于1/3到1/2之间（*Essence*，1）。

正如我们已指出的那样，这类不确定性可能使一个简单的威胁变得太大以至于对于威胁者来说是不可接受的。我们将考察不确定性的一种特定的形式即美国缺乏有关苏联真实动机的知识，并且对其效应进行正式分析。对于所有的其他形式的不确定性，类似结论也成立。

重新考虑图14—1中的博弈。假设苏联从撤出和对抗中获得的支付与原来相反——撤出时为－8，对抗时为－4。在这样的情形里，苏联人持强硬态度。他们宁愿选择核大战带来的毁灭，也不愿意选择屈辱地撤退以及生活在由资本主义美国所主宰的世界中。他们的口号是：“与其生活在星条旗下，还不如死去。”我们在图14—2中画出了此时的博弈树。现在，倘若美国发出威胁，苏联将选择对抗。所以，美国从威胁中获得支付－10；当它不发出威胁并接受古巴导弹的存在时，可获得支付－2。它会“两害相权取其轻”（It takes the lesser of the two evils)。该博弈的这个版本之子博弈完美均衡是苏联“取胜”而美国的威胁失效。

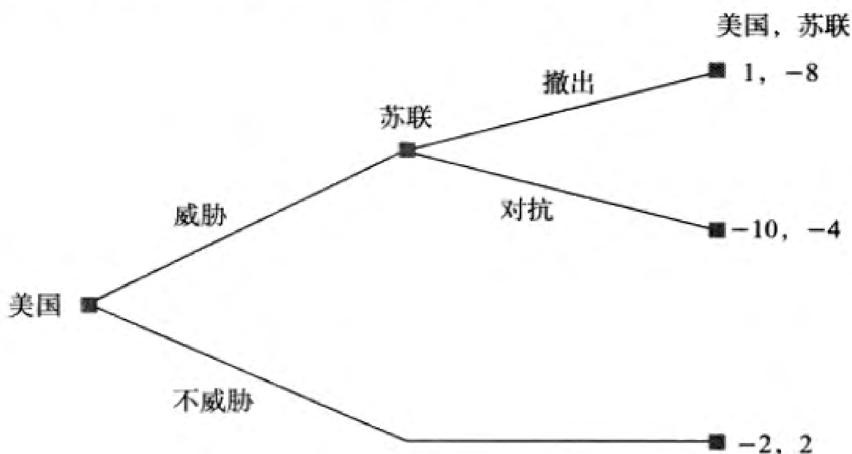


图14—2　与强硬型苏联的博弈

在现实中，当美国人开始选择时，它并不知道苏联是如图14—2所示的那般强硬，还是如图14—1所示的那样较为软弱。美国可以试着估计这两种情况的概率，例如，通过研究苏联过去在不同场合的行为和反应。我们能将肯尼迪所说的封锁引发战争的可能性位于1/3到1/2之间这一概率估计视为苏联是强硬者的概率。由于估计是在一个区间里，我们用一个一般性符号 p 代表这个概率，并研究 p 取不同数值时的结果。

这个更为复杂的博弈的博弈树见图14—3。在博弈开始时，由外部力量（这里标写为“自然”）先决定苏联的类型。在自然选择的上面一枝，苏联是强硬的。这导向其上方的决策结，此时，美国人决定是否发出威胁且博弈树的其余部分与图14—2中的是一样的。在自然选择的下面一枝，苏联是软弱的。这引向其下方的决策结，此时，美国决定是否发出威胁，并且博弈树的其余部分与图14—1中

的一样。但是，美国人并不知道他们自己是在哪一个决策结上。因此，两个美国决策结被包含在一个“信息集”中。这就意味着美国在这个信息集中的不同决策结上不能选择不同的行动，如只在苏联是软弱的时发出威胁等。它在同一个信息集中的不同决策结上只能选择同样的行动，如在两个决策结上都选择威胁或者不威胁。它必须按照真实博弈中位于某一结点的概率来做出决策——也就是说，通过计算两种行动的期望支付来做出决策。

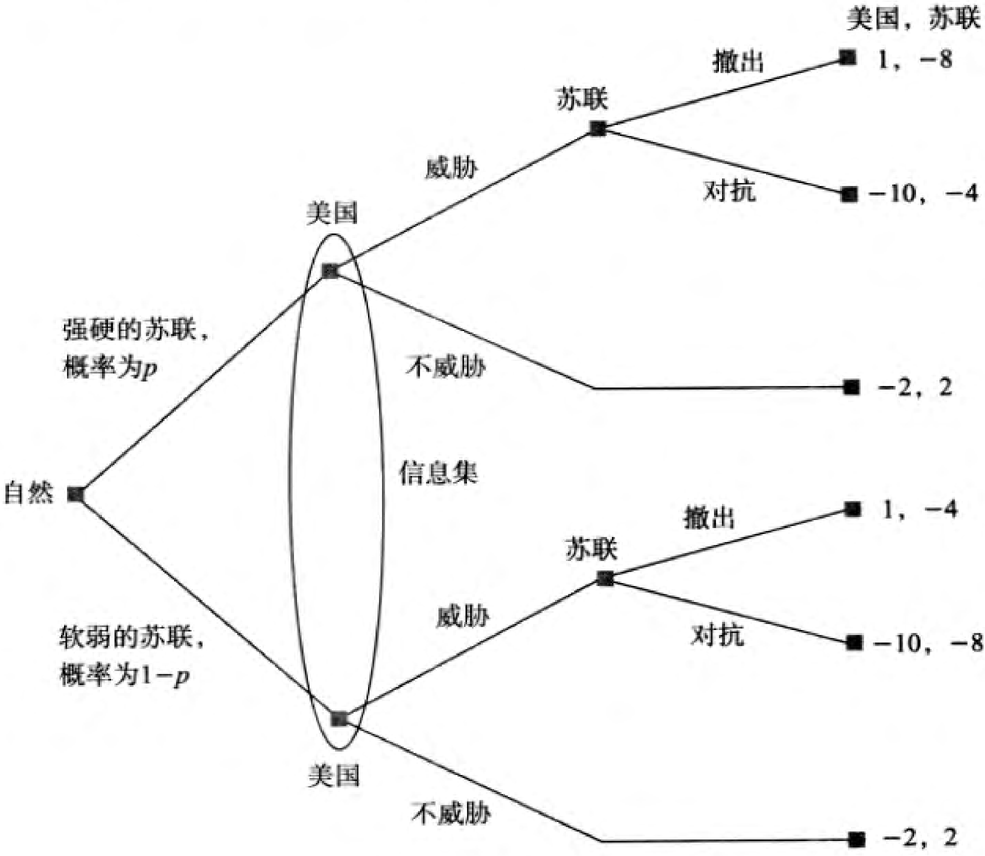


图 14—3　不知道苏联支付下的威胁

当然，苏联是知道自己的类型的。故我们能在博弈结束处展开逆向分析。在上面一枝，强硬的苏联将对抗美国的威胁；在下面一枝，软弱的苏联将在面临威胁时撤回导弹。因此，美国可以预见到，当博弈沿着上面一枝（概率为 p）进行时，选择“威胁”将使其获得支付 -10；而当博弈沿着下面一枝（概率为 $1-p$）进行时，选择“威胁”将使其获支付 1。故美国选择“威胁”的期望支付为

$$-10p+(1-p)\cdot 1=1-11p$$

若美国选择“不威胁”，它将在任一路径上获支付 -2，故其期望支付也为 -2。比较一下两种选择的期望支付，我们看到美国将在 $1-11p>-2$ 或 $11p<3$ 或 $p<3/11$ 时选择“威胁”。

若威胁一定会发挥作用，美国人不会顾及苏联人对抗时他们的支付有多么糟糕，无论是 -10 或更小的负数。但是，苏联人持强硬态度并且因此而对抗威胁

的风险会让美国有可能只得到－10 的支付。仅当苏联人是强硬者的概率 p 充分小的时候，美国人才会发觉发出威胁是可以接受的。因此，给定我们选定的参数值，p 的上限 3/11 也是美国人能容忍的 p 的上限。若我们选择不同的参数，我们将得到不同的上限。例如，若我们将核战争对美国的支付定为－100，则 p 的上限就仅为 3/101。但是，如果威胁因失误而成为现实的概率超过了某个临介点，一般说来，大的威胁将因其结果过于严重而无法兑现。

在本例中，肯尼迪的估计是 p 位于 1/3 到 1/2 之间。不幸的是，该区间的下限 1/3 正好大于我们的上限 3/11——它是美国人能接受的风险上限。因此，这样一种赤裸裸的简单的威胁——“如果你们选择对抗，则将遭遇一场核大战”，实在是太大了，也太冒险了，对于美国人来说代价也太大了。

14.4 概率威胁

如果无条件的战争威胁太大了以至于不能够令人接受，并且也找不到其他的相对较小的威胁，你可通过这样一种方式将威胁变小一些，即创造一种当对方选择对抗时将发生非常可怕的结果的可能性而非必然性。然而，这并非是说你是在事实发生之后才决定是否采取灾难性行动。如果你有着这样一种自由，即你可以避开这种可怕的结果，你的对手将会了解这一点或假定如此，从而威胁一开始就是不可置信的了。你必须放弃某些行动的自由并做出一个可置信的承诺。在此情形下，你必须运用概率机制。

当做出一个简单威胁时，一个参与人对另一个参与人说：“如果你不顺从于我的话，一个对于你来说是十分可怕的灾难就会必然发生。当然，这种灾难对我也是十分可怕的，但我的威胁是可置信的，因为我很重视我的声誉”。当做出一个**概率威胁**（probabilistic threat）时，一个参与人对另一个参与人说：“倘若你不顺从于我的话，一件对于你来说十分可怕的事可能会发生。当然，它对我来说也是一场灾难，但稍后我将无力减少它发生的风险了。”

打个比方，战争发生的概率威胁好比是俄罗斯轮盘赌（在这里真是恰当的比喻）。你把一粒子弹装入左轮手枪的某个弹膛里，然后旋转转轮。子弹就像会引发对双方都有损害的一场核大战的引爆物。当你扣动扳机时，你不知道位于射击道上的弹膛是否已装弹。如果是，你会希望你没有扣动扳机，但此时已太迟了。在扣动扳机之前，若你知道那个弹膛中已装弹，你将不会扣动扳机（即这一行动的后果是极其严重的）。但是，当你知道只有 1/6 的可能性装有子弹时，你是愿意扣动扳机的——此时威胁发生的概率减少到 1/6，现在到了可容忍的那一点上了。

边缘政策（冒险主义）是对这种类型的恰当的风险的一种创造和控制。它包含表面上看来相互不一致的两方面。一方面，你必须让局势在一定程度上失控，以至于你没有充分的自由度在启动了可怕行动之后又去抑制它的发展，从而你的威胁是可置信的。另一方面，你必须对行动发生的风险有足够大的控制力，使你

能将风险控制在不会变得太大以至于你的威胁过于耗费的范围内。这种“受控的失控”(controlled lack of control) 看起来是难以实现的，事实上也的确如此。我们将在 14.5 节中讨论其操作技巧。这里仅提示一下：判断的千差万别、信息的分散，以及执行命令的难度使得一个简单威胁变得极具风险，正是因为如此，制造战争的风险变得可能，并因此而使得边缘政策（冒险主义）变得可置信。实际上的困难不是如何失去控制，而是如何在一种受控的范围内使其失控。

我们先将注意力集中于边缘政策（冒险主义）博弈的机制上。为此，我们将图 14—3 所示博弈稍做改变，见图 14—4。这里，我们假设美国采取另一种威胁，即当苏联选择对抗时，战争将以概率 q 发生，从而 $(1-q)$ 是美国放弃战争并接受苏联在古巴部署导弹的事实的概率。

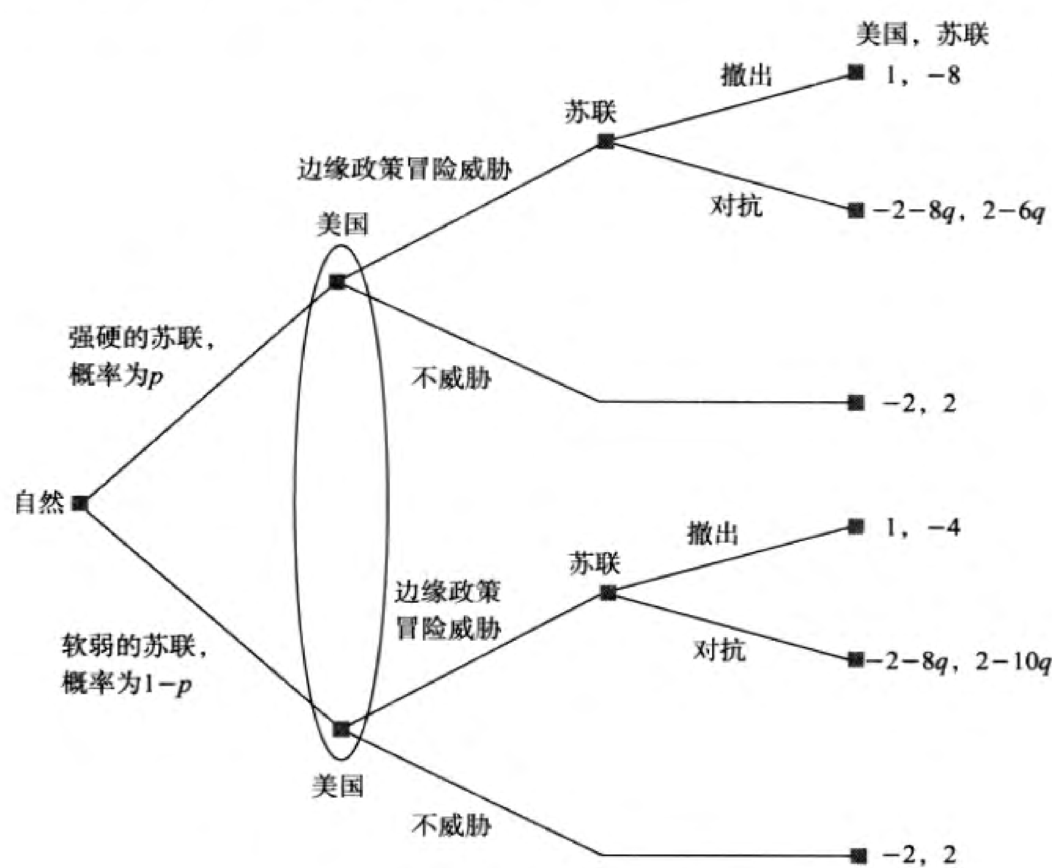


图 14—4　危机的边缘政策（冒险主义）模型

请记住，当博弈进行到苏联对抗美国的时候，后者是没有选择的。俄罗斯轮盘赌已设定概率 q，以此概率将决定在弹膛中是否有子弹（即核大战是否发生）。

因此，若苏联对抗边缘政策（冒险主义）威胁，没有人会准确知道最终将出现的结果和支付，但它知道战争发生的概率 q，也会由此计算期望支付。对于美国来说，结果是以概率 q 得到支付 -10，以概率 $(1-q)$ 得到支付 -2，故期望支付为

$$-10q-2(1-q)=-2-8q$$

对于苏联，期望支付依赖于它是强硬的还是软弱的（它知道自己的类型）。若苏联是强硬的，它从开战中得到-4，以概率q发生；在美国放弃战争时得到支付2，发生概率为$(1-q)$。此时苏联的期望支付为$-4q+2(1-q)=2-6q$。当它选择撤出时，其支付为-8，无论q为0到1之间的任何数值——显然这是最糟糕的。故强硬的苏联将对抗边缘政策（冒险主义）威胁。

如果苏联是软弱的，则有不同的计算。但推理和前面一样，我们知道苏联将从对抗中得到期望支付$-8q+2(1-q)=2-10q$，且若撤出，则肯定得到-4的支付。对于苏联，若$-4>2-10q$，或$10q>6$或$q>0.6$，撤出是较好的选择。故美国的边缘政策（冒险主义）必须含有最少60%发生战争的概率；否则边缘政策（冒险主义）威胁不会令苏联就范，即使它是软弱的。我们称概率q的这个下限为**有效条件**（effectiveness condition）。

将图14—4中的美国边缘政策（冒险主义）和苏联对抗时的期望支付与图14—3中的简单威胁模型对比，可发现，后者现在可被视为图14—4中的一般性边缘政策（冒险主义）威胁模型对应于$q=1$的特别情形的一种特例。

我们可以用通常的方法对图14—4中的博弈求解如下。已知沿上边路径苏联作为强硬者将对抗美国，并且沿下边路径软弱的苏联将在有效条件成立时顺从美国的要求。如果该条件不成立，则两种类型的苏联都将对抗美国；故后者最好不要发出这种威胁。所以，让我们继续假定软弱的苏联将选择顺从，我们看看美国的选择。根本问题是，美国的威胁的风险有多大以及多大的风险是美国可承受的。

倘若美国发出了威胁，它就必须冒着以概率p遭遇到强硬的苏联的对抗的风险。故正如前面所计算的，此时美国的期望支付为$(-2-8q)$。美国人遇到软弱型苏联的概率为$(1-p)$。我们假定他们将顺从，则美国的支付为1。所以，美国从假定可以使软弱型苏联就范（顺从）的概率威胁中得到的期望支付为

$$(-2-8q)\times p+1\times(1-p)=-8pq-3p+1$$

如果美国人不发出威胁，他们得到-2。故美国人发出威胁的条件为

$$-8pq-3p+1>-2$$

或

$$q<\frac{3}{8}\frac{(1-p)}{p}=0.375(1-p)/p$$

也就是说，发生核战争的概率要足够地小，使得它满足该不等式，否则美国根本就不会发出威胁。我们称q的这个上限为**可接受条件**（acceptability condition）。注意，在美国可接受的最大的q的公式中出现了p；苏联对抗的可能性愈大，美国可接受的核战争出现的风险就愈小。

倘若概率威胁有效，它将既满足有效条件又满足可接受条件。我们能够通过图14—5决定战争发生的概率的适当水平。横轴为苏联为强硬型的概率p，纵轴是当它对抗美国的威胁从而导致战争发生的概率q。水平线$q=0.6$给出了有效条件（下限）；要使苏联为软弱型时威胁有效，则与威胁相关的组合

(p，q）必须位于这条水平线的上方。曲线 $q=0.375\ (1-p)\ /p$ 给出了可接受条件（上限）；假定苏联为软弱型时威胁有效，要想让美国人接受战争的风险，与威胁相关的组合（p，q）要位于该曲线的下方。因此，一个有效且可接受的威胁将位于这两条曲线之间的某个区域内，位于它们在 $p=0.38$，$q=0.6$ 的交点的左上方（图 14—5 中的浅灰色阴影区域）。

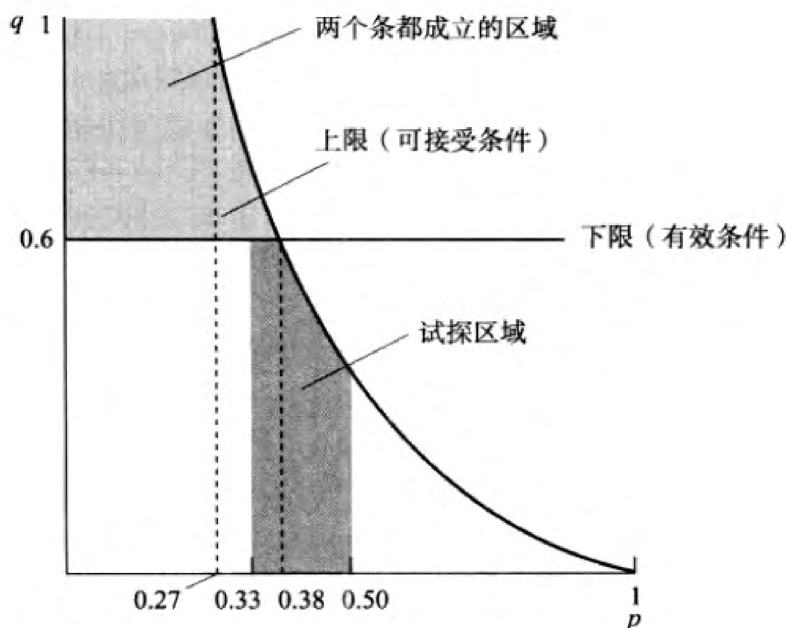


图 14—5　成功的边缘政策（冒险主义）博弈条件

当 $p=0.27$ 时，曲线达到 $q=1$。对于小于这个数值的 p，灾难性的威胁（战争爆发的可能性）对于美国是可接受的，并能恫吓住苏联。这印证了我们在 14.3 节中的分析。

对于从 0.27 到 0.38 内的 p 值，以 $q=1$ 爆发战争的威胁将（p，q）置于可接受条件的右端，并且太大以至于对于美国人来说是不能接受的。但可以找到小一点的威胁。对于这个区域的 p 值，存在对于美国人来说可接受的足够小的 q 值，同时这些 q 值又足够大，可以使软弱型苏联就范。边缘政策（冒险主义）（使用概率威胁）在这种情形可发挥其作用，而一种简单的灾难性威胁将过于冒险了。

如果 p 超过了 0.38，则不存在同时满足两个条件的 q 值。若苏联对抗的概率大于 0.38，则足以迫使软弱型苏联就范的足够大的威胁（$q\geqslant 0.6$）将带来太大的风险，以至于对于美国人来说是不能接受的。$p\geqslant 0.38$ 时，美国人就不可能通过边缘政策（冒险主义）为自己带来任何好处了。

14.5　边缘政策（冒险主义）博弈的操作

如果肯尼迪能准确估计出苏联是强硬者的概率 p 的值，并自信有能力控制因

军事封锁而引发战争的风险 q，则他就能计算和实施他的最优策略。正如我们在 14.3 节中所看到的，若 $p<0.27$，肯尼迪能接受必然发生战争的灾难性威胁（当然，即使在这样的场合，他也偏爱于使用最小的有效威胁，即 $q=0.6$）。

如果 p 位于 0.27 到 0.38 之间，则他就不得不使用边缘政策（冒险主义）。这样一种威胁带有灾难性结果的风险 $0.6<q<0.375\frac{(1-p)}{p}$，且肯尼迪仍然偏爱于这个区域的最小值，即 $q=0.6$。若 $p>0.38$，则他就只有放弃威胁了。

在实际的操作中，肯尼迪不会准确知道 p 的数值，他仅仅能估计它位于 1/3 到 1/2 之间。类似地，他不能准确知道可接受条件中的 q 的临界点的精确位置。对于苏联在各种情形下的支付的数值，如 -8（战争）和 -4（顺从），肯尼迪仅仅能做出一些估计。最后，他可能并不能对由他的边缘政策（冒险主义）带来的风险进行十分精确的控制。所有的这些不确定性使得在控制局势和实施边缘政策（冒险主义）时必须小心翼翼。

假定肯尼迪认为 $p=0.35$，并发出了一个由某种行动所支持的威胁，这一行动将带来战争的风险为 $q=0.65$。该风险大于有效条件的下限 0.6。可接受条件的上限为 $0.375\times(1-0.35)/0.35=0.7$，并且风险 $q=0.65$ 是小于这个上限的。于是，根据肯尼迪的计算，这个风险同时满足两个条件——有效性和可接受性。然而，若假定肯尼迪犯了错误，比如，若他不知道李梅将军实际上可能违背了他的命令并采取了过激行动入侵苏联领空，则 q 实际上高于肯尼迪所设想的值，例如，q 可能等于 0.8，这对于肯尼迪来说风险太大。或者，假定 p 实际上为 0.4，则肯尼迪将认为 $q=0.65$ 也太大了。又或者，肯尼迪的专家顾问们错误估计了苏联人的支付。如果他们将撤出带来的羞辱视为 -5 而非 -4，则有效条件的阈限实际上将为 $q=0.7$，并且肯尼迪以 $q=0.65$ 发出威胁将出错。

肯尼迪所知道的是有效条件和可接受条件曲线的一般形状，如图 14—5 所示。他并不知道 p 的精确数值。因此，他无法准确知道为了保证两个条件都成立应选择什么样的 q 值。实际上，他甚至不知道对于未知的 p 值来说这样的区域是否存在：它可以大于或小于将 p 轴分为两个区域的边界值 0.38。并且，他不能十分精确地确定 q。所以，即使他知道 p，也不能凭自己接受战争风险的意愿去行事。

在这样的模糊环境下，信息不准确，控制不完全，风险大，肯尼迪该怎么办呢？他只有用试探（explore）的办法去寻找苏联和他自己承受风险的能力的边界。在开始试探时，不能在 q 太大的情形下进行。相反，肯尼迪必须从“下面”开始试探边界线。他必须在十分安全的情况下开始试探，并且逐渐地增大风险以便看看“谁将首先眨眼”（Who blinks first）。这就是边缘政策（冒险主义）在实践中的操作过程。

我们用图 14—5 来说明这一过程。看看图中的深灰色阴影区域。它的左右两边的边界 $p=1/3$ 和 $p=1/2$，分别对应于肯尼迪估计的 p 值的区间端点。较低的边界是横轴（$q=0$），较高的边界是由两段曲线组成的。对于 $p<0.38$，该段边界对应于有效条件；对于 $p>0.38$，对应于可接受条件。回忆一下，肯尼迪不知道这些边界的精确位置，但必须从下向上摸索着找到它们。因此，深灰色阴影区

域就是他必须开始进行试探的位置所在。

假定肯尼迪从一个十分安全的行动开始，如q近似等于0.01（1%）。在古巴导弹危机案例中，我们将其看成是他的电视演说，这一演说声明要求苏联的船只立即停下来接受检查。此时，组合点（p，q）位于邻近于深灰色阴影区域的底部边界的位置。肯尼迪不会精确地知道它位于何处，因为他不知道p的精确值。但是，有一点是肯定的，即在此点威胁是十分安全的但也是无效的。因此，肯尼迪开始逐步升级，即他将点（p，q）垂直向上移动。这就是停船检查的实际开始。如果此时仍然是安全但无效的，他就将风险再提高一些。这就是将有关轰炸古巴的计划的信息泄漏一些出去。

随着按此方式不断进行试探，他最终会碰到图14—5中深灰色阴影区域的一个边界，并且到底碰到哪一个边界取决于p的数值。下面两种情形必然会出现。其中一种情形是威胁会变得很大，并足以威胁到苏联。如果p值小于其临界值——这里是0.38，这种情形就出现了。在图中，我们可发现这种情形发生在从深灰色阴影区域移出进入威胁既是可接受的又是有效的区域时。此时，苏联选择退让而肯尼迪获得胜利。另一种情形是威胁对于美国人来说过于冒险了——这在$p>0.38$时发生。肯尼迪的这种方式的试探将他推向可接受条件之上。此时肯尼迪决定退让，并且赫鲁晓夫取胜。我们再次指出：由于肯尼迪不知道p的准确数值，从而不能在事前知道到底哪一种结果会发生。随着他逐步升级地加大风险，他可能会从苏联的行为表现中获取一些线索，这些线索使他对p的估计会逐渐变得精确一些。最后，他会充分准确地知道他位于哪一个边界的下方，并且因此会充分地预知苏联是否会退让或美国应如何行动。

实际上，当肯尼迪通过战争风险逐渐增大的军事行动选择不断试探时，只要灾难性冲突没有发生，就存在两种可能的结果。因此，还存在第三种可能，即在一方认识到已到了风险承受的极限并采取妥协行动之前战争就发生了。持续提升一种很糟糕的结果的风险使得边缘政策（冒险主义）成为一种十分敏感且危险的策略。

因此在实践中，边缘政策（冒险主义）就是一种**相互损害风险的逐步升级**(gradual escalation of the risk of mutual harm)。它可以被生动地想象成**择时躲避的懦夫博弈**（chicken in real time)。在第4章对懦夫博弈进行的分析中，我们简单地假定每个参与人有两种选择：径直向前或者转向。在现实中，选择通常是一个时机问题。两车相向疾驶，每个参与人都可在任何时刻选择转向。当两车相距较远时，转向可以确保安全。随着两人之间的距离愈来愈近，他俩面临的风险不断增大，因为相撞的可能性愈来愈大了。当相距十分近时，即使转向也无法避免相撞了。当他俩继续向对方驶去时，每一参与人都在试探对方愿意承受这种风险的极限，并且也许同时也在寻找自己的极限。第一个达到极限的人将率先转向。但是，总存在这样的风险，即二人到达各自极限的时间都太长，最后他们相距太近以致即使选择转向，仍无法避免相撞。

现在，我们来看为何在古巴导弹危机中，那种使其不能用二人博弈来分析的特征却正好可使边缘政策（冒险主义）更易于实施操作。

相对于立即展开一场核大战而言，军事封锁是一种威胁力度较小的行动。但

是，一旦肯尼迪开始了军事封锁，它的运行、逐步升级以及其他一些方面的变化就不完全在他的控制之中了。故肯尼迪并不是对赫鲁晓夫说："如果你对抗（跨过战争边缘），我就将毫不留情地和必然地发动一场核战争，它将毁灭我们两个国家的人民。"相反，他是说："军事封锁的车轮已经启动起来了，并在聚集自己的能量。你对抗我们的时间愈长，对抗的行为愈多，就愈可能导致局势失控。或者我受到的政治压力将达到这样的一点，它使我退却下来，或者也可能某些鹰派将发动攻击。倘若这一风险不能得到控制，我将不能阻止核战争的爆发，无论对此我会感到多么内疚。现在，只有你顺从我的命令并从古巴撤出导弹，才能缓和紧张的局势，将'嘀嗒'作响的炸弹引信拆除。"

我们相信这一视角比建立在简单威胁基础上的大多数分析能给出对危机更好和更深入的理解。它告诉我们为何战争的风险在所有讨论中扮演着重要的角色。它也使阿里森提出的有关双方的官僚程序及内部分歧的强有力的论述变成了这个图景中的一个组成部分：这些特征让双方的最高级参与人可置信地对下属失去一些控制，即玩起边缘政策（冒险主义）博弈。

还有一个重要条件需要讨论。在第 10 章中，我们看到每一个威胁都有一个与之相关的潜在的承诺，即倘若你的对手顺从你的愿望，就不会出现灾难性的结果。对于边缘政策也同样是这样。倘若在你提升风险水平时，你的对手顺从了你，你应该能立即排除灾难发生的可能。如果随着你增大战争的风险，你的对手屈服了，你必须能"逆向操作"（go into reverse）——立即开始减小风险并将其最终排除。否则，对手从屈服中将一无所获。这在古巴导弹危机中是一个问题。如果苏联人害怕肯尼迪不能控制像李梅将军这一类鹰派人士（"我们今天应该已占领了那里并将他们干掉了"），他们即使选择屈服也将一无所获。

重新强调和加以总结一下，边缘政策是一种将你的对手和你自己置于灾难发生的可能性逐步增大的风险中的一种策略。灾难结果实际发生与否并不全在威胁者的控制之中。

照此看来，边缘政策博弈随处可见。在大多数的冲突中——例如，公司与劳工联盟间的谈判、夫妻间争执、父母与子女的争吵，以及总统与国会之间的角力，一方往往不能准确了解对方的目标和能力。因此，大多数威胁都伴随着犯错误的风险，并且每一个威胁都必然含有一种边缘政策的要素。我们希望我们已经给予了你某些关于这种策略的理解，并且我们也让你对它所带来的风险留下了深刻的印象。在个人生活以及职业生涯里，你会在许多场合里不得不面临或者是由你自己实施边缘政策。请慎用边缘政策，要对其作用和风险有清晰的了解。

为了帮助你谨慎使用边缘政策，我们现在概括一下古巴导弹危机里获取的重要经验，并将之应用于这样一个例子中——一个劳工联盟领袖正在酝酿发起一次旨在增加工资的罢工，但并不清楚该罢工是否会导致工厂倒闭。

第一，开始时要寻找小的和安全的威胁。你的第一步不是立即走出去；应先安排一个持续几天或几周的高层会议，同时与企业主进行谈判。

第二，逐步增加风险。公开或私下声称要罢工，让工人们群情激奋起来。这将诱使管理层相信劳工们接受其当前的低工资的可能愈来愈小了。如果可能，制造一些小的意外事件，如只持续几天的罢工或局部的上街游行。

第三，随着这个过程的持续，从管理层的行动中解读出一些信号，找出企业是否有足够的潜在利润负担劳工们较高的工资要求。

第四，保持对局势发展足够的控制力，即保持引导你的伙伴们批准你与管理层达成的协议的权力，否则管理层会认为即使满足了你的要求也不能减小风险。

14.6 总　结

在某些情形下，实施威胁的过程中存在着犯错误的风险，这就要求用尽量小的威胁。当一个大的威胁不能通过其他方式减小时，则可以通过降低实施威胁的可能性达到减小威胁的目的。概率威胁的策略式应用，就是让你和你的对手都面对一场风险不断增加的灾难，此即"边缘政策"。

边缘政策要求参与人有所保留地放弃对博弈结果的控制。你必须要发出一个具有适当风险水平的威胁——它既是足够大的，能有效迫使你的对手就范；又是足够小的，使得自己也能接受。为了达到这一目的，你必须通过逐步升级相互损害的风险来确定双方的风险承受水平。

1962 年的古巴导弹危机，就肯尼迪总统而言，是运用边缘政策的一个案例。将此危机作为一个以美国对古巴进行军事封锁建立置信度的简单威胁的例子进行分析是不全面的。更恰当的分析是将局势中存在的许多复杂性和不确定性纳入考虑范围，而且需要考虑到简单威胁存在着风险过大的可能性。因为真实的危机涉及大量的政界和军方人士，肯尼迪通过军事封锁，逐步让意外事件和紧张状态升级，做到了"受控的失控"，最后赫鲁晓夫在逐渐增大的核战争风险面前屈服了。

关键术语

可接受条件（acceptability condition）

择时躲避的懦夫博弈（chicken in real time）

有效条件（effectiveness condition）

相互损害风险的逐步升级（gradual escalation of the risk of mutual harm）

概率威胁（probabilistic threat）

习　题

1. 在正文中，我们指出当（每一种类型的）苏联对抗美国的威胁时，美国的支付是－10；图 14—3 标示出了所有这些支付。现在假定该支付事实上是－12 而不是－10。

(a) 在类似图 14—4 那样的博弈树中标示出变化后的支付。

(b) 利用 (a) 所得到的博弈树中的支付，找出美苏边缘政策博弈的这个版本的有效条件。

(c) 利用 (a) 中的支付，找出这个博弈的可接受条件。

(d) 画一个类似于图 14—5 的图，说明在 (a) 和 (b) 中获得的有效条件和可接受条件。

(e) 对于什么样的 p 值，即苏联是强硬者的概率，纯威胁 ($q=1$) 是可接受的？对于什么样的 p 值，纯威胁是不可接受的但边缘政策仍然是可能的？

(f) 肯尼迪认为 p 位于 1/3 到 1/2 之间。如果这一估计是正确的，你对该博弈的这个版本的分析可以得出一个有效且可接受的概率威胁是存在的吗？用这个例子去解释博弈论理论家关于参与人支付的假定是如何对基于理论模型的预测产生重大影响的。

2. 考虑这样一个博弈，它在一个劳工联盟与一家公司之间进行。公司雇用该劳工联盟的成员。劳工联盟能够以罢工相威胁（或不罢工）来迫使公司满足它的工资和利益要求。当面临一个罢工的威胁时，公司能够选择满足劳工联盟的要求或者对抗其罢工威胁。然而，劳工联盟在它决定是否要发出威胁时不知道公司的利润潜力，它不知道公司是否有充分的利润潜力满足其要求（并且公司在这方面的强硬陈词是不可信的）。自然决定公司是否有利润潜力，企业不具有相应利润潜力的概率为 p。支付结构如下：

(ⅰ) 当劳工联盟不发出威胁时，劳工联盟获得支付 0（不管公司的盈利能力如何）；并且当公司具有利润潜力时公司得到的支付是 100，而当公司不具有利润潜力时公司得到的支付是 10。一个温和的劳工联盟会给公司留下较多的利润，如果存在着利润的话。

(ⅱ) 如果劳工联盟威胁要罢工并且公司退却的话，劳工联盟得到 50（不管公司利润潜力的概率如何）；并且当公司具有利润潜力时公司得到的支付是 50，否则其支付为 −40。

(ⅲ) 当劳工联盟威胁要罢工并且公司对抗劳工联盟威胁的话，劳工联盟一定会罢工且得到 −100（不管公司的盈利能力如何）；并且当公司具有利润潜力时公司得到的支付是 −100，而当公司不具有利润潜力时公司得到的支付是 −10。对于一个具有利润潜力的公司来说，对抗会带来巨大的损失，但对于一个不具有利润潜力的公司来说，损失没有那么大。

(a) 假设劳工联盟将采用纯威胁要罢工，除非公司满足劳工联盟的要求。此时会发生什么呢？

(b) 假设劳工联盟造成这样一种局势，其中在公司对抗其威胁之后以概率 $q<1$ 存在着其要进行罢工的某些风险。这种风险来自于劳工联盟的领导人在维持其成员步调一致上的欠完善的能力。为这个博弈画一个类似于图 14—4 的博弈树。

(c) 假设劳工联盟采用边缘政策，以某个未知的概率 q 威胁要罢工，除非公司满足劳工联盟的要求。此时会发生什么呢？

(d) 导出这个博弈的有效条件和可接受条件，并且确定 p 和 q 的数值，并分

析在什么情况下劳工联盟能够采用一种纯威胁、边缘政策，或者是根本就不威胁。

3. 许多电影情节都演绎过边缘政策。从这个视角分析下列描述。双方所面临的风险是什么？在这种边缘政策威胁的实施过程中它们是如何增大的？
   (a) 在 1980 年的电影《上帝必定疯狂》（*The Gods Must Be Crazy*）里的一个情节中，一个被逮捕的试图暗杀一位非洲国家总统的反叛者正在被审讯。他被蒙住双眼背对着一架直升机敞开着的门站立着，在直升机螺旋桨的噪音中，一位官员大声地询问他："谁是你的领导？哪里是你们的藏身之处？"这个男人没有回答，官员就把他推出门外。在下一个画面里，我们看到这架直升机，尽管它的发动机仍然在运行，实际上是在地面上，而且这个男人仰面朝天地跌落了 6 英尺。那位官员出现在门边并且笑着说道："下一次，它会稍微高一点。"
   (b) 在 1988 年的电影《简单计划》（*A Simple Plan*）中，有两兄弟在搬走他们在一架失事飞机上发现的 440 万美元的赎身金。在许多饶有趣味的命运曲折之后，幸存的劫掠者汉克发现自己在与一位联邦调查局的人商谈。这个特工怀疑但不能证明汉克拥有一些丢失的钱，在告知汉克关于这笔巨款的原委后，还告诉他联邦调查局掌握了原来那些赎身金大约十分之一的钞票的序列号。这个特工对汉克说的最后一句话是："现在，事情很简单，就只是等待数字出现了。你不可能带走 100 美元的钞票而最终不给某些人留下印象。"
   (c) 在 1941 年的经典电影《马耳他猎鹰》（*The Maltese Falcon*）中，英雄萨姆·斯佩德［Sam Spade，汉弗莱·博加特（Humphrey Bogart）］是唯一知道镶有宝石饰钉价值连城的猎鹰图在什么地方的人，而恶棍卡斯大帕·古特曼［Caspar Gutman，悉尼·格林斯特里特（Sydney Greenstreet）］正在威胁要使用酷刑让他说出来。斯佩德指出，除非有死亡的威胁，酷刑是没有用的，而且古特曼是不敢杀死他的，否则，秘密就会随他而去了；所以他根本就不害怕酷刑。古特曼回答道："那是一种姿态，先生，那需要彼此间有精妙的判断，因为，正如你知道的那样，先生，人们在行动过火时往往忘记他们的利益所在，最后难免感情用事。"

4. 在这道习题里，我们提供几个成功运用边缘政策的例子，其中"成功"是指双方都做到了相互可接受。对于每一个例子，（ⅰ）识别出参与人的利益；（ⅱ）描述出每种情形下内在的不确定性的本质；（ⅲ）给出参与人用来使得灾难风险逐步升级的策略；（ⅳ）讨论这样的策略是否是好的策略；（ⅴ）（选做）如果你能够做到的话，为本章里的内容建立一个小小的数学模型。在每一种情形里，我们都提供一些读物让你开始；你应该通过运用图书馆的资源和诸如 Lexis-Nexis 这样的互联网来找到许多的东西。
   (a) 从 1986 年开始并且导致 1994 年形成世界贸易组织的乌拉圭回合国际贸易谈判。读物：John H. Jackson，*The World Trading System*，2nd ed.（Cambridge：MIT Press，1997），pp. 44－49 and Chaps. 12 and 13。
   (b) 1978 年以色列与埃及之间的《戴维营协议》。读物：William B. Quandt，

Camp David: *Peacemaking and Politics* (Washington, DC: Brookings Institution, 1986)。

(c) 1989 年到 1994 年南非种族隔离政权与非洲国民大会之间为建立一个多数人规则的新的宪法体制所进行的谈判。读物：Allister Sparks, *Tomorrow Is Another County* (New York: Hill and Wang, 1995)。

5. 在这道习题中，我们提供几个未能成功运用边缘政策的例子。这里，当对双方都有害的结果（灾难）发生时，边缘政策就被认为是"不成功的"。在下列情形中回答习题 4 中提出的问题：

(a) 在美国，1995 年克林顿总统与共和党控制的国会之间的预算案冲突。读物：Sheldon Wolin, "Democracy and Counterrevolution," *Nation*, April 22, 1996; David Bowermaster, "Meet the Mavericks," *U. S. News and world Report*, December 25, 1995/January 1, 1996; "A Flight that Never Seems to End," *Economist*, December 16, 1995。

(b) 1991 年到 1998 年的卡特彼勒公司罢工。读物："The Caterpillar Strike: Not Over Till It's Over," *Economist*, February 28, 1998; "Caterpillar's Comeback," *Economist*, June 20. 1998; Aaron Bernstein, " Why Workers Still Hold a Weak Hand," *Business week*, March 2, 1998。

【注释】

［1］两个十分优秀的基于理论的案例研究的例子是 Pankaj Ghemawat 的《商业博弈：案例与模型》(*Games Businesses play*: *Cases and Models*, Cambridge: MIT Press, 1997)，以及 Robert H. Bates, Avner Greif, Margaret Levi, Jean-Laurent Rosenthal 和 Barry Weingastd 的《讲故事》(*Analytic Narratives*, Princeton: Princeton University Press, 1998)。

［2］我们的资料来源包括 Robert Smith Thompson 的《十月的导弹》(*The Missiles of October*, New York: Simon&Schuster, 1992)；James G. Blight 和 David A. Welch 的《在深渊的边缘：美国和苏联重新审视古巴导弹危机》(*On the Brink*: *Americans and Soviets Reexamine the Cuban Missile Crisis*, New York: Hill and Wang, 1989)；Richard Reeves 的《肯尼迪总统：权力的历程》(*President Kennedy*: *Profile of Power*, New York: Simon&Schuster, 1993)；Donald Kagan 的《论战争的起源与维持和平》(*On the Origins of War and the Preservation of Peace*, New York: Doubleday, 1995)；Aleksandr Fursenko 和 Timothy Naftali 的《赌博的地狱：古巴导弹危机秘史》(*One Hell of a Gamble*: *The Secret History of the Cuban Missile Crisis*, New York: Norton, 1997)；以及最后的、最近的和最为直接的《肯尼迪的录音带：在古巴导弹危机期间的白宫》(*The Kennedy Tapes*: *Inside the White House During the Cuban Missile Crisis*, ed. Ernest R. May and Philip D. Zelikow, Cambridge: Harvard University Press, 1997)。Graham T. Allison 的《决策的本质：解释古巴导弹危机》(*Essence of Decision*: *Explaining the Cuban Missile Crisis*, Boston: little Brown, 1971) 是重要的，不仅仅由于其叙述的故事，而且是因为其分析和理解。我们与他在某些重要的方面存在着观点上的分歧，但是我们受益于他的洞察力。我们依据并扩充了 Avinash Dixit 与 Barry Nalebuff 的《策略思维》(*Thinking Strategically*, New York: Norton, 1991) 第 8 章中的思想。当我们引用这些资料去

证明特定的论点时，均在文章中做了标注。标注是从书的题目中选择的关键词以及相关页码。

[3] 对于那些不能获得第一手信息的或者想寻找关于这次导弹危机的戏剧性故事和细节的初步介绍的人，我们推荐电影《十三天》（*Thirteen Days*，2000，New Line Cinema）。

[4] 参见 Richard Rhodes，*Dark Sun*：*The Making of the Hydrogen Bomb*（New York：Simon &Schuster，1995），pp. 573－575。李梅其人，以其极端观点和不停地咀嚼没有点燃的雪茄烟而闻名，被认为是 1963 年拍摄的电影《奇爱博士》中的 Jack D. Ripper 将军的原型。这位将军命令他的轰炸机飞到苏联去发动一场突然袭击。

[5] 在 2002 年 10 月于古巴哈瓦那举行的纪念导弹危机 40 周年的一次会议上，这一事件被公开了。参见 Kevin Sullivan，"40 Years After Missile Crisis，Players Swap Stories in Cuba，" *Washington Post*，p. A28。Vadim Orlov 是苏联潜艇部队的一员，他认识那位拒绝发射鱼雷的军官 Vasili Arkhipov。Vasili Arkhipov 逝世于 1999 年。

第 15 章　策略与投票

选举中的投票乍一看来似乎处于博弈论的视野之外，但是，记得 2000 年美国总统选举的人应该不会对在选举和投票中有着许多策略运用之发现感到惊奇吧。第三方选举候选人拉尔夫·纳德（Ralph Nader）的出现对于选举结果有着许多的策略意味。一个被称作“策略投票公民”的组织花钱在主要的报纸上做广告向普通公众解释投票策略，并且有分析家认为纳德的候选人资格可能就是导致小布什最终当选的因素。

事实表明，选举中除了决定喜欢谁并投票给他外，还有其他一些因素在起作用。因为存在着许多不同类型的投票程序，它们被用来决定选举中的赢家，所以就存在着许多不同的选举结果。某一特定的程序可能带来你所喜欢的一种结果，此时，你就可以运用你的策略思维能力观察到该程序的使用。或者你会发现为一个特定的选举所采用的程序是可以加以操纵的，选民能够通过策略性地不实地表达他们的偏好而影响选举结果。如果其他选民以这种方式策略性地投票，则你最好也这样干。对于选举的策略性的这种理解，有助于你在这种情形里确定你的最优行动步骤。

在接下来的几节中，我们首先考察一下各种类型的投票规则和程序，然后讨论特定情况下会出现的某些非直观的或是悖论性的结果。我们将考虑一些用来评判不同选举方式绩效的标准。我们还会着重介绍投票者的策略性行为和操纵结果的范围。最后，我们将会介绍著名的中位投票者定理（median voter theorem）的两种不同版本——具有离散性策略和具有连续性策略的两人零和博弈。

15.1 投票规则和程序

许多选举程序都可以用来帮助人们从备选项（也可以说，候选人或者议题）中做出选择。在只有三个备选项的情形中，选举设计就会变得十分有趣而复杂。在本节中我们将对三种投票或投票统计方法的各种各样的程序进行描述。可能的投票程序的数量是巨大无比的，如允许选举建立在程序组合的基础上，我们在这里所提供的简单分类就可大大地扩大。在经济学和政治科学中有相当可观的文献就正好是处理这种问题的。我们不打算给出一个关于此类文献的详尽的综述，而只是稍加提及。如果你有兴趣，我们建议你查阅有关此主题更详细的文献。[1]

15.1.1 二元法

投票统计方法可以按照由选民在任何给定时间所考虑的保留权及候选人的数量进行分类。**二元法**（Binary Methods）要求选民每次只能在两个候选人中做选择。在刚好只存在两个候选人的选举中，选票的统计方法就是非常著名的**绝对多数规则**（majority rule）。这一规则简单地规定拥有过半数选票者获胜。当有超过两个的候选人时，可以使用**配对投票**（pairwise voting）——这一方法要求一直重复以二元法进行选举。配对规程是**多阶段的**（multistage），选民必须不断地确定谁是赢家。

有这样一个配对规程，其中每个候选人在遵循绝对多数规则的循环选举中都被提名反对其他每一个候选人，这一规程被称作**孔多塞方法**（Condorcet method），这一理论由 18 世纪法国理论家孔多塞最先提出。他认为在一系列的一对一竞赛中击败每一个对手的候选人，应该获得整场选举的胜利；这样一个候选人，现在都被称为**孔多塞赢家**（Condorcet winner）。其他配对规程会得出一个“得分”，例如**柯普兰指数**（Copeland index），这一指数被用来统计在循环竞赛中所取得的输赢成绩。世界杯足球锦标赛第一轮确定一组当中的哪个队可以参加下一轮的比赛时使用的就是一种柯普兰指数。[2]

另外一个有名的配对规程适用于存在三个候选人时。这一规程就是**修正程序**（amendment procedure）。当对某项法案采用投票这一形式进行表决时，美国国会就要求使用该程序。国会通过某一法案之前，任一法案的修正案都必须首先赢得对法案原始版本的胜利，然后再和现行法案配对，议员们通过投票来确定是否采用第一轮中胜利的法案的版本。绝对多数规则被用于决定胜利者。修正程序被用于考察任何三个候选者，它首先在第一回合的选举中配对两个候选者，然后在第二轮的投票中把第三个候选者和第一轮中的胜利者进行比较。

15.1.2 相对多数方法

相对多数方法（plurative methods）允许选民同时考虑三个甚至更多的选择。一种相对多数投票方法根据候选人在选民选票上的位置对候选人进行评分，这种分数在计算总得分时使用；这种投票方法就是著名的**定位方法**（positional methods）。常见的**相对多数规则**（plurali-ty rule）是定位方法的一种特殊形式，其中每个投票者为他所最偏爱的选择投出单一的一张选票，当加总票数时，该选择就被赋予 1 分；获得最多的投票数（或分数）者取胜。要注意的是那些相对多数方法里的胜者并不需要获得绝对多数或超过 51%的票数。因此，举个例子来说，在 1994 年美国缅因州的州长选举中，独立候选人安格斯·金仅仅获得了 36%的票数就当选为州长，他的民主党、共和党，以及绿党对手们各自获得了 34%、23%以及 6%的选票。定位方法的另一种特殊形式是**反相对多数方法**（antiplurality method），要求投票者投票反对某个候选人，或等价地说是投票支持除了某个人外的其他人。为了进行计算，被投反对票的候选人被分配－1 分，或者除那个人之外的所有其他候选人得到 1 分而被投反对票的人接受 0 分。

最为有名的一种定位方法是**博尔达计数法**（Borda count），它是以让-查尔斯·德·博尔达（Jean-Charles de Borda）的名字来命名的，博尔达是一位乡下小伙子，与孔多塞是同时代的人。博尔达将这一程序描述为对相对多数规则的一种改进。博尔达计数法要求投票者把选举中所有可能的候选人进行排名，并把这种排名写在他们的选票上，并根据每个候选人在选票上的位置来给他们评分。在三人选举中，在选票上排名第一的候选人获得 3 分，排名第二的候选人获得 2 分，排名第三的候选人获得 1 分。在收集了选票之后，每个候选人的得分被加总，分数最高的赢得选举。博尔达计数法在许多与体育运动有关的比赛中被采用，包括职业棒球赛赛扬奖（Cy Young Award）以及大学橄榄球锦标赛。

可以通过改变对候选人的评分规则来简单设计许多其他的定位方法，而这种评分是基于候选人在选民选票上的位置进行的。评分时可能会采用这样的方法，即给予选票上排名最高的候选人比其他候选人高得多的得分，比如，在有三个候选人的选举中，给排名第一者记 5 分，而给排名第二和第三者分别记 2 分和 1 分。在有众多候选人（比如 8 个）的选举中，位于选票上前两名的候选者会被区别对待，分别获得 10 分和 9 分，而其余的候选者则都得到 6 分或更少。

可取代定位相对多数方法的另一种方法是近些年来发明的**赞成投票**（approval voting）方法，该方法允许投票人给他们所“赞成”的候选人中的每一位投出单一的一票。[3] 与定位方法不同的是，赞成投票方法并不会基于候选人在选票上的位置而对于不同候选人有所不同。相反，所有的赞成投票都被平等地对待，且收到最多赞成票者获胜。如果选举允许不止一个候选人胜出（例如选举学校董事会成员），那么赞成票数的一个阈限水平会被提前设置好，那些获得超过所要求的最小赞成票数的候选人将被选中。这种方法的倡导者指出，与处于两头位置的候选人相比，这种方法更有利于处于中间位置的候选人；反对者则声称，易受

欺骗的投票者有可能在他们的选票上给予太多的“鼓励性”赞成票，从而选出某个他们并不想支持的新候选人。尽管存在这些不一致的意见，一些职业性社团已经采用赞成投票方法去选举他们的官员，一些州已经使用或正在考虑用这种方法去进行公共选举。

15.1.3 混合方法

一些多阶段投票程序将相对多数方法与二元法加以组合以形成**混合方法**(mixed methods)。例如，**过半数决胜**（majority runoff）程序就是一种二阶段方法，用来把多种可能性减少为二元决策。在第一阶段选举中，投票者投票给他们最偏好的候选人，并且投票情况被记录下来。如果某个候选人在第一个阶段得到了大部分的选票，他将赢得选举。然而，如果没有候选人得票过半数，则在第二个阶段选举中，两个最受支持的候选人进行比赛，由过半数规则选出胜者。法国总统选举使用的就是过半数决胜竞选程序。假如有三个或四个强有力的候选人在第一回合中分享选票，那么这一方法可以产生意想不到的结果。例如，在2002年春天，极右派候选人勒庞（Le Pen）在第一轮的总统选举中遥遥领先于法国社会党总理若斯潘（Jospin），从而进入了第二轮。这个结果在法国人民中引起了惊恐和不安，30%的公民甚至在选举中都没有打算去投票，其中有些人把第一轮当做表达其对左翼候选人的偏好的一种机会。勒庞进入决胜局在政坛上激起了轩然大波，虽然在最后他没有登上总统宝座。

另一种混合程序涉及连续多轮的投票。投票者在每一轮（rounds）中对众多候选人进行投票，表现最差的候选人将被剔除。投票者在下一轮中继续对剩下的候选人进行投票。剔除的过程一直持续到仅剩下两个候选人为止，此时，通过二元法和过半数决胜程序来决定最后的胜者。这种多轮程序被用于选择奥林匹克运动会的主办地点。

人们可以通过让选民在第一轮所投的选票上表达出他们的偏好顺序来消除对连续多轮投票的需要，**单一可传递投票**（single transferable vote）方法在后面的回合中可用来统计投票数。使用单一可传递投票方法时，每个投票者在开始的选票上表达出他们对所有候选人的偏好顺序。如果没有候选人在第一轮投票中获得过半数的选票，排名最后的候选人则要被淘汰，所有将该候选人列在第一位的选票就被转移给选票上排名第二的候选人；随着更多的候选人被剔除，类似的重新分配发生在后面的选举轮中，直到某个获得过半数选票的胜者出现为止。2002年3月在旧金山所进行的市政官员选举中使用的就是这种方法。这一方法更普遍地被称为**即时决胜**（instant runoff）方法。

单一可传递投票方法，在选举中有时与**比例代表**（proportional representation）方法联合使用。例如一个州的选民中55%的人是共和党人，25%的人是民主党人以及20%的人是独立候选人，比例代表方法意味着它将产生出一个反映该州选民党派组成的代表性团体。这种结果与相对多数方法形成强烈的对比，相对多数方法选举出的可能全部是共和党人（假设在每个地区不同党派选民的比例

与整个州中不同党派选民的比例一致）。获得一定份额选票的候选人都被选出，且其他没有达到一定份额的候选人则被淘汰，这都取决于投票程序中具体的规定。按照投票者对候选人的排序，将被淘汰的候选人的选票转移。这种程序继续进行下去直到来自每个党派中的适当数量的候选人被选出为止。

很清楚，在投票统计方法的选择上，存在着相当大的可进行策略性思维的空间。即使统计规则已被选定，策略也仍然是很重要的。我们将在下一节中考察一些与规则制定和议程有关的问题。进一步，投票者的策略性行为——常常被称为**策略性投票或不实的偏好策略性陈述**（strategic voting or strategic misrepresentation of preferences），能够在任何规则下影响选举结果。在本章后面的章节中，我们将会看到这一点。

15.2 投票悖论

即使人们是根据他们个人真实的偏好进行投票的，投票者偏好以及投票程序上的特定条件都可能产生出不可思议的结果。另外，选举结果关键性地依赖于票数统计中所用的程序类型。本节将介绍一些最为有名的不可思议的结果——所谓的投票悖论——以及一些这样的例子，即在投票者偏好不变和不存在策略性投票时，不同的票数统计方法是如何改变选举结果的。

15.2.1 孔多塞悖论

孔多塞悖论（Condorcet paradox）是最著名和最重要的投票悖论之一。[4]就像前面所提及的，孔多塞方法所产生的胜者是在两人配对竞赛轮次中每一轮均获得过半数选票的那一个候选人。当没有孔多塞赢家从这一过程中产生时悖论就出现了。

为了说明这一悖论，我们构造一个例子，其中三个投票者使用孔多塞方法对三个可选择的结果进行投票。考虑三位城市议员（“左”、“中”和“右”），他们被要求对三个可供选择的福利政策的偏好进行排序。其中一个政策是要扩大当前可获得的福利利益（称其为慷慨的，或者记为 G），另一个则是减少可获得的福利利益（称其为减少的，或者记为 D），再一个则为维持现状（称其为一般的，或者记为 A）。然后要求他们就每一对政策进行投票，以建立一种议会排序，或**社会排序**（social ranking）。这种排序意味着对议会作为一个整体怎样判断可能的福利制度的优劣进行描述。

假设议员“左”倾向于尽可能地维持高福利，而议员“中”则最愿意维持现状，但也关注城市的预算情况，且最不愿意去增加福利。最后，议员“右”最愿意减少福利但倾向于现有福利有所增加，他预期增加福利将会引起一场严重的预算危机，它会剧烈地扭转公众对于福利的意见，从而使得一种更加持久的低福利

状态由之产生，而这种现状可无限期地持续下去。我们用表 15—1 来说明这些偏好排序，其中弯曲的“大于”号“≻”，被用来表示一个备选项优于另一个备选项。(从技术上来说，“≻”指一种二元关系。)

有了这些偏好，如果 G 与 A 配对，则 G 将取胜。在接下来的配对中，是 A 与 D 配对，A 赢。在最后 G 与 D 配对中，比分再一次是 2 比 1，这时有利于 D。所以，如果议员对可选择的政策对进行投票，过半数的是 G 胜过 A，A 胜过 D，以及 D 又胜过 G。没有一项政策会同时过半数胜过另外两项。群体所表现的偏好是循环的：G≻A≻D≻G。

表 15—1　　议员对福利政策的偏好

左	中	右
G≻A≻D	A≻D≻G	D≻G≻A

这种偏好循环是**不可传递偏好顺序**（intransitive ordering of preferences）的一个例子。理性概念通常意味着个人偏好顺序是**可传递的**（transitive）。如果一个人被给定有 A、B、C 三种选择，并且你知道他喜欢 A 超过 B，喜欢 B 超过 C，则可传递意味着他喜欢 A 超过 C。(该术语来自数学中的数字大小的传递性，例如，如果 3>2 并且 2>1，那么我们知道有 3>1。) 与从我们的城市议会例子中导出的那个社会偏好顺序不一样，一种可传递的偏好顺序并不会循环。所以，我们说，这样一种顺序是不可传递的。

注意，所有的议员对三种可供选择的福利政策都有可传递的偏好，但是议会却没有。这就是孔多塞悖论，即使所有个人的偏好顺序都可以传递，这并不能保证由孔多塞投票程序所产生的社会偏好顺序是可传递的。这种结果对于公众服务者以及普通公众来说，有着非常深远的意义。它使“公共利益”这一基本概念存在问题，因为这些利益并不能简单地加以定义，甚至可能是不存在的。我们城市议会在福利政策上并没有任何定义完好的群体偏好集。其教训是，社会、机构或者是其他大型的人民群体并不总是可以被解读成像个人那样行动。

孔多塞悖论甚至可能会更加一般地出现。不能确保由任何规范的群体投票过程产生的社会偏好顺序会因个人偏好顺序可传递而可传递。然而，一些估计表明，在大型群体对大量候选人做出选择时，这种悖论最有可能出现。较小的群体考虑较少数量的候选人时，更容易对这些候选人有类似的偏好，在这种情况下，这种悖论则很少出现。[5] 事实上，在我们的例子中，这种悖论会产生是因为议会不仅在哪个候选人最好上而且还在哪个候选人最糟糕上完全存在分歧。越小的群体中这种结果越不太可能出现。

15.2.2 议程悖论

我们所考虑的第二个悖论也涉及一种二元投票程序，但是它考虑了该程序中

备选项的顺序。在设有一名对存在三个备选项的选举有决定投票顺序的权力的委员会主席的议会中，该主席在最终结果上有巨大的权力。最后结果的实际决定权掌握在主席手上，事实上，他可以利用产生于某些个人偏好集合的不可传递的社会偏好顺序，通过选择一种适当的议程，以他需要的任何方式去操纵选举结果。

再次考虑城市议员“左”、“中”、“右”，他们必须要在 G、A 和 D 福利政策中做出决策。议员们对备选项的偏好如表 15—1 所示。现在让我们假设其中一位议员被市长任命为议会主席，且主席有权决定先对哪两个福利政策进行投票，哪个政策可以在第二轮投票中再和在第一轮投票中获胜的政策对决。在给定议员偏好及偏好顺序的共同知识的情况下，主席可以获得任何他想要的结果。例如，如果议员“左”被选为主席，他就可以安排 A 和 D 在第一轮中对决而胜者再和 G 在第二轮中对决，以此得到 G 取胜的结果。**议程悖论**（agenda paradox）指的就是这样一种结果，即任何最终排序都可通过选择适当的程序得到。

根据表 15—1 中所表达的偏好集，我们不仅得到“孔多塞悖论”下的不可传递的社会偏好顺序，而且还可以得到这样的一种结果，即二元程序下任何一个备选项成为赢家都是可能的。在这样一个案例中，决定结果的唯一因素就是议程的顺序。这种结果意味着，在这里议程的设定才是真正的博弈，由于主席能够设定议程，主席的委派或选举才是策略性行为的关键所在。此时，如同其他许多的策略性情形一样，看起来是博弈的（在这个案例中，是福利政策的选择），其实并不是真正的博弈；反而那些稍早参加策略性游戏（决定谁是主席）和最后根据在最终选举中设定的偏好进行投票的人，才是真正参与博弈的人。

然而，我们在前面对议程设定者的力量的说明中假定，在第一轮中，投票者在两个备选项中（A 和 D）做出选择，这种选择仅仅基于他们对这两个备选项的偏好，并不考虑程序的最后结果。这样一种行为被称为**诚实投票**（sincere voting），实际上，短视的或非策略性投票可能会是一个更好的名称。如果议员“中”是一个策略性博弈参与人，他应该会意识到，如果他在第一轮中投 D 的票（即使在那个阶段所提名的两个备选项中他更喜欢 A），则 D 将在第一轮中胜出，并在第二轮中，将会由于得到议员“右”的支持而击败 G。议员“中”更希望 D 而不是 G 成为最后的结果，所以他应该做逆推分析且在第一轮中进行策略性投票。但如果其他每一位投票者也都进行策略性投票的话，他会这样做吗？我们将会在 15.4 节中考察策略性投票博弈并寻找其均衡。

15.2.3 颠倒悖论

定位方法同样也可以导致自相矛盾的结果。例如，当公开给选民的候选人名单发生变化时，博尔达计数法可带来**颠倒悖论**（reversal paradox）。选票已全部提交之后剔除其中一个候选人时，这种悖论会产生于至少有四个候选人的选举中，使得有必要进行重新计算。

假设有四个候选人竞争专门颁给退休的大联盟棒球投手的特别纪念赛扬奖。候选人分别是史蒂夫·卡尔顿（SC）、桑迪·科法克斯（SK）、罗宾·罗伯茨

(RR) 以及汤姆·西弗 (TS)。七位著名的体育专栏作家应邀在他们的投票卡上对这些投手进行排名。在每张投票卡上排名第一的候选人会获得 4 分；第二名、第三名和第四名的候选人将会依次获得 3 分、2 分和 1 分。

通过七位作家的投票，产生了对候选投手的三种不同的偏好顺序。这些偏好顺序以及拥有每一种偏好顺序的作家的数量，见表 15—2。当所有选票被计分后，西弗得到 (2×3)+(3×2)+(2×4)=20 分；科法克斯得到 (2×4)+(3×3)+(2×1)=19 分；卡尔顿得到 (2×1)+(3×4)+(2×2)=18 分；还有罗伯茨得到 (2×2)+(3×1)+(2×3)=13 分。所以，西弗赢得选举，接下来则依次为科法克斯、卡尔顿和罗伯茨。

现在假定罗伯茨被发现是不具有参赛资格的，因为他实际上从未赢得过赛扬奖，他的职业顶峰所出现的年份早于该奖设立的 1956 年。这一发现要求在排除罗伯茨的得票后重新计算得分。现在在每一张投票卡上排名第一者获得的是 3 分，而排名第二和第三者分别得到的是 2 分和 1 分。例如，从具有偏好顺序 1 的运动专栏作家那里得到的投票，现在分别给予科法克斯和西弗 3 分和 2 分，而不是在第一次统计时的 4 分和 3 分；在最后的位置，这些投票同样给卡尔顿 1 分。

表 15—2　体育专栏作家对投手们的偏好

顺序 1(2 个投票者)	顺序 2(3 个投票者)	顺序 3(2 个投票者)
科法克斯＞西弗＞罗伯茨＞卡尔顿	卡尔顿＞科法克斯＞西弗＞罗伯茨	西弗＞罗伯茨＞卡尔顿＞科法克斯

重新计分后，卡尔顿就获得 15 分，科法克斯获得 14 分，而西弗获得了 13 分。这种结果将输家变成了赢家，把第一次的选举排名颠倒过来了。专栏作家并没有改变其偏好顺序，但却出现了这种奇怪的结果。两次选举中唯一的不同就是候选人的数目。

15.2.4 改变投票方式就改变了投票结果

就像前面的讨论里所说明的那样，不同的投票规则下选举结果很可能是不同的。举例子来说，假设有 100 名投票者，根据他们对三个候选人 (A、B 和 C) 所持有的不同的偏好将他们分成三组。表 15—3 显示了三种不同的偏好。这三个候选人中的任意一个都能赢得选举，取决于所采用的投票加总方法。

在简单相对多数规则下，候选人 A 以 40%的选票获胜，尽管有 60%的投票者将其置于三个候选人中的最后一位。候选人 A 的支持者们显然将偏爱于这种类型的选举。如果他们有权力选择投票方法的话，那么简单相对多数规则这一看似“公平”的程序将会使得 A 赢得选举，尽管大多数人强烈地不喜欢该候选人。

表 15—3 对候选人的群体偏好

第一组（40 个投票者）	第二组（25 个投票者）	第三组（35 个投票者）
A≻B≻C	B≻C≻A	C≻B≻A

然而，采用博尔达计数法，将会产生一种不同的结果。博尔达规则中，最受欢迎的候选人可得 3 分，中间者得 2 分，最不受欢迎的得 1 分。候选人 A 在 40 张选票中获得第一，在 60 张选票里获得第三，总计为 40(3)＋60(1)＝180 分。候选人 B 在 25 张选票里获得第一，在 75 张选票里获得第二，总计为 25(3)＋75(2)＝225 分。候选人 C 在 35 张选票里获得第一，在 25 张选票里获得第二，40 张选票里获得第三，总计为 35(3)＋25(2)＋40(1)＝195 分。在这种程序下，候选人 B 获胜，C 居第二，A 排在最后。在反相对多数投票规则下，候选人 B 也会赢，其中投票者将投票支持除其最不喜欢的候选人之外的所有候选人。

那么对于候选人 C 又怎么样呢？如果采用过半数决胜规则或者即时决胜规则，他将会赢得选举。在任何一种方法中，候选人 A 和 C 在第一轮各赢得 40 张和 35 张选票，得以幸存下来进入决胜局。过半数决胜规则会要求选民再次考虑候选人 A 和 C 并进行投票；即时决胜规则会剔除掉候选人 B，然后把候选人 B 的得票（是来自第二组的投票）分配给其第二偏好的候选人 C。那么，由于候选人 A 是 100 个选民中 60 个都排在最不喜欢的位置上的候选人，候选人 C 将会在决胜局中以 60 比 40 取胜。

另一个例子是挑选夏季奥林匹克运动会主办地的投票，同样说明了不同的程序是如何导致不同的结果的。挑选奥林匹克主办地的实际投票程序，是一种混合方法，由有剔除的多轮相对多数规则和最终的过半数决胜规则投票组成。参与投票的每个国家的代表投出一票，每一轮投票中获得最低总票数的候选城市被剔除。过半数决胜规则被用于在最后的两个候选城市之间进行选择。在对 1996 年和 2000 年奥运会主办地的投票中，都存在一个城市在开始的大多数轮中一直是领先的——1996 年，雅典（希腊）在前几轮中都是领先的，而在 2000 年，北京一开始就领先。在这两个案例中，这些相对多数赢家在其他城市被剔除后，没有得到多少额外的支持，而且在每一场选举中领先者都没有最终取胜。倒数第二轮及之前的每一轮采用的都是相对多数方法，但是在最后的剔除轮之后，雅典看到其所需要的额外票数被投向了亚特兰大，而北京眼瞧着它们被投向了悉尼。奥林匹克主办地的选举程序一旦被改变了，之前的一切就都失去了任何意义。

15.3 关于悖论的一般性思路

上一节对各种各样的投票悖论的讨论说明投票方法可能由于许多瑕疵而产生不寻常的、未曾预料到的，甚至是不公平的结果。另外，这也引起我们的疑问：

是否有某种投票规则能满足一定的规定性条件，包括传递性，并且也是最为“公平的”——也就是说，最为准确地表达出选民们的偏好？肯尼斯·阿罗（Kenneth Arrow）的**不可能性定理**（impossibility theorem）告诉我们，对这个问题的答案是：没有。[6]

虽然阿罗不可能性定理的技术内容使得进行完全的证明超出了我们的讨论范围，但是该定理的含义却是很容易把握住的。阿罗指出不存在满足他所提出的所有下列 6 项关键性原则或者准则的偏好加总方法：

（1）社会或群体排序必须能将所有候选人进行排序（完备的）。

（2）是可传递的。

（3）它要满足被称为“正反应”（positive responsiveness）或帕累托性质（Pareto property）的条件。给定两个候选人 A 和 B，如果选民们一致地表现出偏好 A 胜于偏好 B，则加总后的排序应将 A 放在 B 之前。

（4）排序不能是由独立于社会中个体成员的偏好的外部因素（如习俗）所强加的。

（5）它一定不能由独裁决定——任何单个的选民不能够决定群体的偏好顺序。

（6）它应独立于无关候选人，也就是说，候选人集合的变化（加入或退出）不应该引起对未受影响的候选人的排序的变化。

通常，该定理的前四个条件会被精简掉，而把重点放在同时满足最后两个条件的困难上。简化形式的定理指出，偏好加总排序不可能在没有独裁的情况下独立于无关候选者（IIA）。[7]

你能立即看到前面所考虑的某些投票方法并不能满足阿罗不可能性定理的所有原则。例如，条件 IIA 就被单一可传递投票程序以及博尔达计数法所违反。然而，博尔达程序是非独裁的和一致的，并且它满足帕累托性质。我们所考虑的所有的其他投票规则都满足 IIA，但却违反了其他原则中的某一个。

阿罗不可能性定理激起了对改变潜在假设后其结论的稳健性的广泛研究。经济学家、政治科学家以及数学家们都在研究减少准则数目或将阿罗原则放松至最小从而找到一种可在不牺牲核心原则情况下满足这些准则的程序的方式。他们的努力大部分都是不成功的。大多数的经济学家和政治理论家们现在都接受这样一个观点，即当选择一种投票或偏好的加总方式时，某种形式的妥协是必要的。这里有一些著名的例子，每个都代表了一个特定领域——政治科学、经济学和数学——的方法。

15.3.1 布莱克条件

正如在 15.2 节中所表明的那样，配对投票程序在社会偏好顺序的传递性上并不满足阿罗不可能性定理中的条件，甚至当所有个人偏好顺序都可传递时也是如此。克服这一障碍以满足阿罗不可能性定理条件的一种方法是在单个投票者的偏好顺序上施加限制，这一方法同样可以避免孔多塞悖论。这样一种限制条件，被称为**单峰偏好**（single-peaked preferences）的必要条件，是由政治科学家

邓肯·布莱克（Duncan Black）于 19 世纪 40 年代末提出的。[8]布莱克论述群体决策的开创性论文实际上早于阿罗不可能性定理，且表述中也包含了孔多塞悖论的思想，但是投票理论自那时起便显示了它与阿罗工作的相关性；事实上，单峰偏好的必要条件有时就被称作**布莱克条件**（Black's condition）。

对于单峰偏好顺序来说，一定是这样的，被考虑的选项可以沿着某一特定维（例如，与每一政策相联系的支出水平）进行排序。为了说明这一条件，在图 15—1 中我们画出了这样一幅图，水平轴上是特定维（政策），垂直轴上是选民偏好（或支付）的排序。为使单峰偏好成立，每个选民必须有一个单一的最理想或最为偏好的候选人，并且"远离"最偏好点的候选人一定提供稳定的较低支付。图 15—1 中的两个选民"左"先生和"右"先生，在政策主张上有着不同的最理想点，但对每一投票者而言，当政策远离他或她的最理想点时，其支付会稳定地下降。

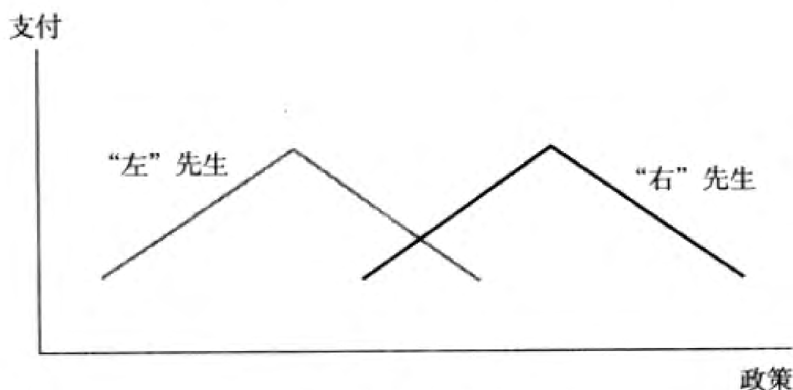


图 15—1　单峰偏好

布莱克指出，如果每个投票者的偏好都是单峰的，那么配对（过半数）投票程序一定会产生出一种可传递的社会偏好顺序。从而，孔多塞悖论得以避免，并且配对投票满足阿罗的可传递条件。

15.3.2　稳健性

有一种另外的选择，它来自于由经济理论家达斯古普塔（Partha Dasgupta）和马斯金（Eric Maskin）所提出的一种更为现代的对于阿罗不可能性定理来说是妥协的方法。[9]他们提出用一种新的被叫做**稳健性**（robustness）的准则来评判投票方法。稳健性是这样测度的，考察一个非独裁的且能满足 IIA 以及帕累托性质的投票程序同时满足其社会偏好顺序的可传递性要求的频率，即，对于多少投票偏好顺序的集合来说这种程序满足可传递性的要求。

通过采用稳健性的准则，可以证明简单绝对多数规则是最为稳健的——也就是说，它是非独裁的，满足 IIA 和帕累托性质，且对于最大可能的投票者偏好顺序集合来说提供了可传递的社会偏好顺序。从稳健性上看，其他的投票程序比绝对多数规则差，包括博尔达计数法和相对多数规则。这种稳健性准则，在制定一种最为通常使用的投票程序作为最好的加总程序的备选方案的能力上，十分

具有吸引力——这种程序最为经常地与民主过程相联系。

15.3.3 强度排序

另外一类理论为避免得出阿罗的负结果，把关注点放在了满足阿罗 IIA 要求的困难上。最近的一种这类理论来自数学家唐纳德·萨里（Donald Saari）。[10] 他建议投票加总方法应更多地利用有关选民偏好的信息，而非仅限于他们对任何一对候选人 X、Y 的排序中包含的信息，可把每个投票者对候选人 X、Y 的偏好强度也考虑进来。这种强度可以通过计算一个投票者置于 X 与 Y 之间的其他候选人 Z、W、V 的数目来加以测量。萨里因此用一种不同的准则取代了 IIA 条件，即用他标记为 IBI（二元独立强度）的准则取代了阿罗的第六条原则，我们将其记为第（6′）条原则：

> （6′）任何两个候选者的社会相对排序可以仅由每个选民在该对中的相对排序和这种排序的强度来加以确定。

这个条件要比 IIA 弱一些，因为它看来是把 IIA 仅仅应用于“不相关”候选人的增添或剔除上，并不改变人们在“相关”候选人之间偏好的强度。经过这种修改，博尔达计数法就能满足修正过的阿罗不可能性定理了，并且它是唯一一种能满足修正过的阿罗不可能性定理的定位投票方法。

萨里把博尔达计数法视为唯一恰当地对投票计数内部平局加以评价的程序，他认为它对于一种好的加总规则来说是应满足的一种基本的准则。平局会以两种方式发生：在投票者偏好中通过**孔多塞意义**（Condorcet terms）或通过**颠倒意义**（reversal terms）。在有三个候选人 A、B、C 的选举中，孔多塞意义是偏好顺序 A≻B≻C，B≻C≻A，及 C≻A≻B。有着这些偏好的三种投票的集合的一次性投票结果表达在逻辑上是相互抵消的，从而形成一个平局。颠倒意义是这样一种偏好顺序，它包括一对候选人位置上的颠倒。在同样的选举中，具有 A≻B≻C 和 B≻A≻C 这样一种偏好顺序的两种投票结果逻辑上可导致在 A 和 B 的成对对决中产生平局。只有博尔达程序才将这种投票结果加总——有着孔多塞意义或者颠倒意义——处理成为平局。就像在前面章节中所看到的那样，虽然博尔达计数法可以导致颠倒悖论，但它仍有许多倡导者。博尔达程序仅当选票收拢后再从中剔除掉某些候选者的时候才产生自相矛盾的结果。因为当对所有候选人集采用唯一的一次投票时，这种结果就可以避免，所以博尔达程序在一些循环中被视为最好的选票加总方法之一而受到青睐。

对于一个好的加总规则应该满足的准则，其他一些研究者提出了不同的建议。它们中的某些准则包括孔多塞准则（Condorcet criterion）（如果存在的话，孔多塞赢家应该可以通过一种投票规则而被选出来）、一致性的准则（在一个包括所有投票者的选举之中当选者应与将全体投票者任意划分成两部分后所举行的两次选举的当选者是同一人），以及无操纵性准则（投票规则不应该在选民中鼓励策略性投票）。我们不可能长篇大论地来考察这些建议中的每一种，但是我们

要在下面两节中讨论选民的策略性操纵行为。

15.4　投票的策略性操纵

我们已考察过的若干投票规则对于选民来说存在着策略性地不实地陈述其偏好的较大余地。在 15.2.2 节中，我们已经分析了议员“中”为淘汰他最不喜欢的候选人，并将他更喜欢的候选人送进第二轮，是如何在第一轮投票中违背自己真实的偏好挑战设定议程的主席“左”的权威的。更为一般地，如果选民可以通过某些行为按照其意愿改变最终的选举结果，那么选民就会投票给之前选举中出现的并非他们所最喜欢的候选人。在本节中，我们将考察策略性投票行为影响选举的诸多方式。

15.4.1　相对多数规则

相对多数规则选举尽管常被许多选民们认为是最为公平的，也仍然提供了策略性行为的机会。例如，在总统选举中，通常有两大主要候选人展开竞争。当其支持度相对接近时，很有可能会有第三个候选人加入竞争并将领先者的选票分走。若第三个参与人的加入真的威胁到领先者赢得这场选举的机会，那么这位后来的进入者就被称为**分肥者**（spoiler）。

分肥者一般被认为没有多大机会赢得整个选举，但其在改变选举结果中所起的作用却是无可争议的。在存在分肥者的选举中，那些最偏好分肥者，其次偏好目前领先的主要候选人，最后才是偏好目前掉尾的候选人的选民，最好策略性地不实地陈述其偏好，以避免他们最不偏好的候选人当选。也就是说，在这种情况下，即使你比较喜欢分肥者，但由于分肥者不大可能获得多数票，因此你应该将票投给目前领先的候选人；将票投给目前领先的候选人可以阻止你最不喜欢的，也就是目前掉尾的候选人当选。[11]

有证据显示，在近几次美国总统大选中——特别是 1992 年和 2000 年的选举中——存在着大量的策略性投票。在 1992 年，罗斯·佩罗（Ross Perot）在第一季的大多数时间里掀起了强大的第三方竞选造势运动，但却在盛夏时退出选举，而在最后又重新参选。由于他再次参选的时间太迟，他被视为当时总统乔治·布什和其民主党挑战者比尔·克林顿之间势均力敌竞争中的分肥者。

人们对佩罗是否分走了共和党选票而有助于克林顿在那一年获胜是存在着争议的。不过，若没有佩罗的参选，克林顿极有可能是无法当选的。一种更加有意思的理论认为，在投票中存在着明显的偏好扭曲。《新闻周刊》（*Newsweek*）上的一份调查声称，如果更多的选民相信佩罗有实力在选举中获胜，他可能最终就会获胜；接受调查的选民中 40%的相对多数说，如果他们认为佩罗会获胜的话，他们都会投佩罗（而不是布什和克林顿）的票。[12]

拉尔夫·纳德（Ralph Nader）在2000年的总统竞选中也在民主党副总统阿尔·戈尔（Al Gore）和共和党挑战者乔治·布什之间扮演过类似的角色。纳德更为关心的是能否获得5%的大众选票而不是能否取得这场总统竞选的胜利，因为如果获得5%的选票，他的绿党就有资格获得联邦竞选基金（federal matching election funds）。但是他的候选人资格使竞争激烈的选举更加紧张。2000年的竞选也因佛罗里达州格外接近的投票结果以及该州的选票之战而更加地令人难以忘却。但是纳德在竞选操作中从戈尔阵营里拉来所需的选票，引起了足够的关注，公众纷纷讨论（在广播电视和各主要城市报纸上）其策略性投票的可能性。特别地，有几个组织（以及许多的网站）提出了旨在使纳德在不会造成戈尔在其任何的关键州丢失选票的情况下获得其所需选票的“选票交换”（vote swapping）计划。在戈尔的关键州（如宾夕法尼亚州、密歇根州和缅因州）中支持纳德的选民们被要求与在那些注定会支持乔治·布什的州（如得克萨斯州或怀俄明州）中的戈尔支持者“交换”他们的选票——密歇根州的纳德支持者投戈尔的票，而其支持纳德的选票在得克萨斯州投。有关这些策略是否灵验的证据是模糊不清的。我们知道的是纳德没有赢得5%的大众选票但戈尔拿走了宾夕法尼亚州、密歇根州和缅因州的所有选票。

在立法机构的选举中，要挑选出很多候选人，第三方在全民普选比例代表制下的表现与在分立选区多数代表制下的表现有很大不同。在不列颠采取的是选区制和相对多数代表制。在过去的50年中，工党和保守党一直是轮流坐庄。自由党虽然在竞选中获得了第三高的支持率，却一直以来是策略性投票的受害者，且因此在议会中得到与其地位不相称的很少的席位。意大利实行的是全民普选制和比例代表制。在此规则下并不需要进行策略性投票，并且即使小的政党也有可能在立法机构中获得不少席位。经常是没有党派拥有明显多数的席位，且小党可通过讨价还价进行结盟而对政策产生影响。

如果在影响国家的政治选择方面非常不具效率的话，一个政党是无法兴旺发达的。因此，我们往往看见在使用相对多数代表制的国家里仅产生两个主要政党，而在那些使用比例代表制的国家里会有好几个政党。政治学家将此现象称为杜弗杰法则（Duverger's law）。

在立法机构里，选区制倾向于产生仅仅两个主要政党——经常出现的情形是其中一个政党拥有明显过半数席位从而政府有较大决定权。但是会存在少数党的利益被忽视的风险——也就是说，会产生“多数人专制”的风险。比例代表制则给予了少数党更好的发出声音的机会。然而，它却会产生出对于权力的没完没了的讨价还价式的角逐和立法僵局。有趣的是，每个国家都觉得其自身的制度表现欠佳并且在考虑要向着另外一种制度规则转变：在不列颠，有很大的呼声要求要实行比例代表制，而意大利则一直都在认真地考虑是否要实行选区制。

15.4.2 配对投票

当你知道你只能使用诸如修正程序那样的配对方法时，你就可以采用你对第

二轮的结果的预测来决定你在首轮中的最优投票策略。在第一轮中，对某一候选人或者政策做出承诺是符合你的利益的，甚至在它并非你最为偏好的选择时也是如此，这样你最不喜欢的选择就不会在第二轮中赢得整个选举了。

在这里，我们回到城市议会的例子中去，其中有一位有权设定议程的议会主席。再一次强调，我们假设整个议会都已知道所有关于三个候选人的偏好顺序。假设最偏好 G 的议员"左"被任命为主席，并将 A 和 D 安排在第一轮中进行投票相互对决，其中的获胜者再于第二轮中迎战 G。若三位议员都严格按照如表 15—1 所示的偏好进行投票的话，A 将会在第一轮投票中击败 D，而 G 会在第二轮投票中赢定 A；主席所最为偏好的结果就会被选出。然而，议员们看来都是一些训练有素的策略家，他们都会预见到最终一轮的投票，并且会运用逆推法决定他们在前面回合中应采取的投票方式。

在刚才所描述的情形里，在选举中，议员"中"所最不偏好的政策将会当选。因此，逆推分析表明，他应该在第一轮投票中进行策略性投票，以影响选举结果。若议员"中"在第一轮投票中将票投给他最偏好的政策，他就会投 A 的票，而 A 会在这一轮中击败 D 并在第二轮投票中输给 G。然而，在第一轮投票中，他也能够策略性地投 D 的票，这样就会将 D 在第一轮中提升到 A 之上。然后，在第二轮投票中，D 会被安排与 G 进行对抗，G 就会输给 D。议员"中"在 A 和 D 上不诚实地陈述其偏好顺序有助于其将赢家从 G 改变为 D。尽管 D 并非他最偏好的结果，但从他的角度看，它也比 G 要好。

如果议员"中"能够肯定在此选举中不存在其他的策略性投票，那么这种策略就很有用。所以我们需要全面分析两轮投票，来验证对于这三个议员而言的纳什均衡策略。我们通过对在该选举中两个同时进行的投票回合使用逆推法，并从两种可能的第二回合竞争——即 A 对 G 或者是 D 对 G——出发，来完成这一工作。

表 15—4 分析了在每一种可能出现的第二轮选举结果。表 15—4（a）中的两个表格表示当 A 赢得第一轮而要与 G 决胜时的获胜政策（并不是参与人的支付）；表 15—4（b）中的表格表示当 D 赢得首轮胜利时的获胜政策。在这两种情形里，议员"左"选择最后结果的行，议员"中"选择列，而右则选择表格（左边的或者右边的）。

你可以知道，每一个议员在每一种第二轮选举中都有一个占优策略。在 A 对 G 的选举中，议员"左"的占优策略是将票投给 G，议员"中"的占优策略是投给 A，而议员"右"的占优策略投给 G；所以 G 会赢得这次选举。如果议员们考虑 D 对 G 的选举，议员"左"的占优策略仍然是投给 G，议员"中"和"右"的占优策略都是投给 D；在这场选举中，D 获胜。快速的验证表明，在这一轮里所有的议员都是按照他们的真实偏好进行投票的。因此这些占优策略都是"投我所偏好的候选人的票"。由于在第二轮投票里不必考虑未来，议员们也只是简单地投在他们自己的偏好顺序中排位较靠前的政策的票。[13]

我们现在可以使用我们对表 15—4 的分析结果来考虑第一轮选举中的最优策略，其中选民们在政策 A 和 D 之间进行选择。因为我们知道下一轮中议员们会如何投票，哪一项政策会获胜，所以我们就能够在表 15—4 的表格中表示出整个选举的结果。

表 15—4　　两种可能的第二轮投票中的选举结果

(a) A 对 G 的选举

"右" 投票：

A

		中	
		A	G
左	A	A	A
	G	A	G

G

		中	
		A	G
左	A	A	G
	G	G	G

(b) D 对 G 的选举

"右" 投票：

D

		中	
		D	G
左	D	D	D
	G	D	G

G

		中	
		D	G
左	D	D	G
	G	G	G

我们以表 15—5 中位于右边表格左上方栏中的 G 为例来说明是如何得到这些结果的。当议员"左"和议员"中"都在第一轮中投给 A 而议员"右"投给 D 时，就得到这个单元格中的结果。因此，A 和 G 就会在第二轮中被配对投票，如表 15—4 所示，G 获得胜利。其他结果也可以用类似的方法得到。

给定表 15—5 中所示的结果，议员"左"（其为主席，并已设定议程）在这一轮中的占优策略是投 A 的票。同样，议员"右"的占优策略是投票给 D。没有任何一个议员在任一轮中颠倒其偏好或进行策略性的投票。然而对于议员"中"来说，在这里即使他对 A 的偏好严格超过对 D 的偏好，投票给 D 也是占优的策略。正如在前面的讨论中所显示的那样，他在第一轮投票中有强烈的动机去颠倒其偏好，并且他是唯一一个进行策略性投票的人。议员"中"的行为改变了选举的结果，使得当选者由 G（不存在策略性投票时的赢家）变成了 D。

表 15—5　基于第一轮投票的选举结果

"右" 投票：

A

		中	
		A	D
左	A	G	G
	D	G	D

D

		中	
		A	D
左	A	G	D
	D	D	D

回忆一下那位主席，议员"左"，他设定议程是为了让他最为偏好的候选人

获胜。然而，选举结果却是他最不偏好的候选人当选了。看起来设定议程的权力根本就没有多大用处。但是，议员“左”会预期到策略性行为。那么，他就可以利用对策略博弈的理解挑选议程。事实上，若他在第一轮中安排D对G，其获胜者再挑战A的话，则纳什均衡结果就会是G，即主席最偏好的结果。在这种议程中，议员“右”为了阻止他最不喜欢的A当选，会在第一轮中颠倒他的偏好，将票投给G，而非D。可以验证，这就是议员“左”的最佳议程设定策略。在整个博弈中设定议程被看做初始的、先于投票的一轮博弈，我们可以预料到在议员“左”是主席时，G会得到采纳。

当我们更加仔细地审视选举的策略性版本中的投票行为时，就能够发现出现了一种有趣的模式。有成对的议员在两轮选举中都“一起”投票（大家投的票都一样）。在原有的议程下，议员“右”和“中”在两轮中都是“一起”投票的，而在建议的议程（第一轮是D对G）下，议员“右”和“左”在两轮中都是“一起”投票。换句话说，两种情形下，在两位议员之间都形成了一种长期持续的联盟。

国会中在许多的场合里都发生过这种类型的策略性投票。一个例子是1956年的联邦学校建设基金的提案。[14]在提交国会就不建立基金而维持现状进行投票之前，众议院对提案进行了修改，要求只为没有种族隔离学校的州提供资助。在国会的投票规则下，首先进行的投票是决定是否接受所谓的鲍威尔修正案，之后再考虑胜出的提案版本。研究这个提案的历史的政治科学家们指出，学校基金的反对者们为了击败原有的提案而在对修正案的投票中策略性地不实陈述了他们的偏好。一个关键性的代表集团开始投票赞成修正案，但是随后在最终的投票中又加入了种族融合反对者的阵营而投票反对整个提案，所以这个法案最终没有通过。该集团的投票记录表明，其中有许多人在其他场合里曾经投票反对种族融合，这意味着他们投票赞成种族融合仅仅是一种暂时的策略性投票行为，并不代表他们对学校种族融合的真实看法。

15.4.3 不完全信息下的策略性投票

前面的分析表明，有时候委员会的成员们有动机去采取策略性的投票方法来阻止他们最不喜欢的候选人赢得选举。我们的例子假定议员们知道可能的偏好顺序以及有多少其他议员具有这样一些偏好。现在假定信息是不完全的：每个议员都知道可能的偏好顺序、他自己的实际偏好顺序，以及其他每个议员具有特定的偏好顺序的概率，但并不知道其他议员中的不同偏好顺序的真实分布。在这种情况下，每一个议员的策略需要依他对此分布的信念以及他对其他人会在多大程度上按照他们自己的真实意愿进行投票的信念的具体情况而定。[15]

举个例子来说，假定我们仍然有一个有三位成员的议会，他们在考虑三种可选择的福利政策。这些政策就是在前面按照议员“左”所设定的（原有的）议程所描述的那种。也就是说，议会在第一轮中考虑政策A与D，再由胜出者与G在第二轮中对决。我们假定仍然存在着三种不同的可能偏好顺序，如表15—1所

示，并且议员们知道这些顺序是全部的可能性。不同的地方在于，没有人确切地知道各个偏好集有多少议员。然而，每一个议员知道他自己的类型以及观察到其他的类型（左，中或右）的概率，设 P_L、P_C及 P_R分别为其观察到议员类型为左、中、右的概率，它们之和为 1。

我们在前面已经看到，在最后一轮的投票中，三个议员的投票都是按照其真实的意愿进行的。我们还看到，议员“左”和“右”在第一轮中的投票也是真实的。这种结果在不完全信息的情形下仍然成立。右类型市议员乐于看到 D 赢得首轮选举；给定这一偏好，议员“右”投票给 D 而不是 A 总会做得一样好（如果其他两个议员也按照同样的方式投票），有时通过这种方式做得更好（如果其他两个议员分别投 D 和 A 的票）。类似地，左类型市议员更乐于看到 A 避免被淘汰而能够在第二轮中与 G 对决；这些议员总是能够做到与其他做法至少一样地好，并且有时会做得更好——通过投票给 A 而不是 D。

问题仅在于中类型议员身上。因为他们并不知道其他议员的类型，而他们又有动机根据某些偏好分布进行策略性地投票——特别是在肯定知道每一种类型只有一个投票者的情形下，他们的行为将依赖于议会中各种投票者类型出现的概率。我们在这里考虑两种极端情形中的一种，其中有一个中类型议员相信其他中类型议员将会按照其真实的意愿进行投票，此时我们将发现一个对称的、纯策略的纳什均衡。在他相信其他中类型议员将策略性地进行投票的情形下，他将采取习题 9 中的投票策略。

为了能对结果进行比较，我们将对中类型议员与可能赢得选举的政策相联系的支付指定特定的数值。中类型议员的偏好是 A≻D≻G。假定，如果 A 取胜，中类型会得到支付 1；如果 G 赢了，中类型得到支付 0；如果 D 赢，中类型会得到某一中间水平的支付，把它记为 u，其中 $0<u<1$。

现在，假定我们的中类型议员在他相信不管其他市议员的类型如何，其他两个议员都会按照其真实意愿进行投票的选举里，必须决定在第一轮中（A 对 D）怎么投票。如果其他两个议员选择的是同一政策，要么是 A 要么是 D，则中类型市议员的投票对于最终结果是不重要的；他在 A 与 D 之间是无差异的。但是，如果其他两个议员将选票分别投给不同的政策，则中类型议员就可以影响最后的选举结果。他的问题是需要决定是否要按其真实的意愿进行投票。

如果其他的两个投票者分别把票投给 A 和 D 并且他们都是按照自己的真实意愿进行投票的，则 D 的得票必然来自于右类型议员。但是 A 的得票既可能来自左类型议员，又可能来自于（按照真实意愿的）中类型议员。如果 A 的得票来自左类型议员，则中类型议员就知道每一种类型中都有一个议员。如果在这种情形下他按照其真实意愿投票给 A 的话，A 就会赢得第一轮但在后来会输给 G；中类型议员的支付将会是 0。如果中类型议员策略性地把票投给 D，D 就击败 A 和 G，且中类型议员的支付为 u。另一方面，如果 A 的得票来自中类型议员，则中类型议员就知道在议会中有两个中类型议员和一个右类型议员但却没有左类型议员。此时，对 A 进行按照真实意愿的投票是有助于 A 赢得首轮投票的，并且在随后的第二轮中 A 也将以 2 比 1 的票数击败 G；中类型议员获得其最高支付 1。如果中类型议员策略性地把票投给 D，D 将会再次赢得两轮投票并且中类型

议员将得到 u。

为了确定中类型议员的最优策略，我们需要对其从按照真实意愿进行投票中所得到的期望支付和从策略投票中得到的期望支付进行比较。按真实意愿投票给 A 时，中类型议员的支付取决于 A 的另一张得票有多大可能是来自左类型或者右类型议员。这些可能性是很容易计算出来的。A 的另一张得票来自左类型议员的可能性就是左类型议员作为剩下来的投票者中的一员的概率，或 $p_L/(p_L+p_C)$；类似地，A 的票来自中类型议员的概率是 $p_C/(p_L+p_C)$。于是中类型议员从真实意愿投票中获得的支付就以概率 $p_L/(p_L+p_C)$ 取 0，以概率 $p_C/(p_L+p_C)$ 取 1，所以期望支付是 $p_C/(p_L+p_C)$。在对 D 进行策略性投票时，不管第三个议员是什么类型，D 将获胜——D 肯定会胜出——因此中类型议员期望支付就恰好是 u。只要 $p_C/(p_L+p_C)>u$，中类型议员最终的决策就是按照其真实意愿进行投票。

注意，中类型议员的决策制定条件在直观上是合理的。如果有更多中类型议员的可能性比有一个左类型议员的可能性更大，则中类型议员就会按照其真实意愿进行投票。仅当他是议会中属于其类型的唯一一个投票者时，对于中类型议员来说进行策略性投票才是有用的。

我们对不完全信息的存在以及其对策略性行为的潜在影响再做两点评论。首先，如果议员的数目 n 大于 3，并且是奇数的话，那么中类型议员从策略性投票中得到的期望支付将仍然是 u，且从按照其真实意愿进行投票中得到期望支付为 $[p_C/(p_L+p_C)]^{(n-1)/2}$。[16] 所以，仅当 $[p_C/(p_L+p_C)]^{(n-1)/2}>u$ 时，中类型议员才会按照其真实意愿进行投票。由于有 $p_C/(p_L+p_C)<1$ 以及 $u>0$，对于足够大的 n 来说，该不等式绝不会成立。这个结果告诉我们在一个足够大的议会中，一种对称的按照真实意愿进行投票的均衡绝不可能持续下去！其次，有关其他议员偏好的不完全信息，为策略性行为创造了额外的机会。在有多于两轮投票的议程中，议员们可以通过他们在前面几轮里的投票来传递有关其类型的信号。剩下几轮就给予了其他议员更新其对概率 p_C、p_L、p_R 的先验信念的机会和在这样的信息基础上采取行动的良机。在仅有两轮的成对投票中，在第一轮里没有时间利用任何获得的信息，因为在最后一轮里对于所有议员来说，按照其真实意愿进行投票都是占优策略。

15.5 与操纵有关的一般性思路

一种投票程序易于受到偏好的策略性不实陈述或者 15.4 节所说明的不同类型的投票者的策略性操纵所影响的程度，是在投票理论家中已经引起广泛注意的另外一个话题。阿罗在他的理论中并不要求非操纵性，但是在他的文章中，他也考虑了这样一种需求如何与阿罗条件相关联的问题。类似地，理论家们已经考察过在各种各样的程序中可以进行操纵的可能性，并且由此对投票方法进行了排序。

经济学家威廉·维克里（William Vickrey），或许因其对拍卖的研究而更为人所知（见第 16 章），做了一些研究选民们的策略性投票行为情况的早期工作。

他指出，满足阿罗 IIA 假设条件的程序，是最不容易受到策略性操纵影响的。他也设定了一些条件，在这些条件下，策略性行为更易发生并取得成功。他特别注意到给定投票方法本身是可操纵的，拥有较少知情选民和较小可选择候选人集合的情形，可能是最容易受到操纵影响的。然而，这个结果意味着，为了使得投票程序满足阿罗条件而弱化 IIA 假设条件会导致更多的操纵性程序。特别地，在 15.3.3 小节里所提到的萨里的 IIA 假设条件的强度排序版本（称为 IBI），可能会允许更多的程序满足阿罗定理的一种修正版本，但是可能同时也会允许更多具有可操纵性的程序去满足这一定理。

正如阿罗在偏好加总的不可能性上的一般性结果一样，在操纵性上的一般性结果是一个负结果。具体地说，**吉巴德—萨特斯韦特定理**（Gibbard-Satterthwaite theorem）表明，如果所要考虑的候选人有三个或更多，唯一可阻止策略性投票的选举程序是独裁：某一个选民被指定为独裁者，并且其偏好就决定了选举结果。[17] 将吉巴德—萨特斯韦特定理与维克里关于 IIA 的讨论结合起来，就可以帮助读者理解为什么阿罗不可能性定理经常被简略为考虑何种程序能够满足非独裁性和 IIA 条件。

最后，某些理论家们指出，不应该根据其满足阿罗条件的能力，而应该根据其鼓励操纵的倾向性来对投票规则进行评估。一种投票规则的相对操纵性可以由关于其他选民偏好的信息数量来决定，这些信息是选民们成功地操纵一场选举所需要的。基于这种准则的某些研究表明，就目前所讨论的程序来说，相对多数规则是最易进行操纵的（也就是说，只需要最少的信息）。接下来按其可操纵性的递减顺序排列的程序分别是赞成投票、博尔达计数法、修正程序、绝对多数规则，以及仿真兔程序（Hare Procedure）（单一的可传递投票）。[18]

按可操纵性的水平来对程序所进行的分类，仅仅取决于操纵一种投票规则所必需的信息数量，而并不是对该信息的充分利用，或者个人或集团是否能易如反掌地进行操纵。实际上，单个选民操纵相对多数规则是相当困难的。

15.6 中位投票者定理

我们在前面所有的章节里讨论的都是存在多个候选人的选举中投票者的行为、策略及其他方面。然而，策略性分析也可应用于这类选举中候选人的行为上。例如，给定选民的特定分布和选民的偏好，候选人就需要决定提出其政治纲领时的最佳策略。当在选举中只有两个候选人，选民们沿着政治光谱按照一种“合理的”方式分布，并且每一个选民都有“合理的”一致性（意味着是单峰）的偏好时，**中位投票者定理**（median voter theorem）告诉我们，两个候选人都会将他们自己置于政治光谱中与中位投票者相同的位置上。**中位投票者**（the median voter）就是在那个分布中的“中间”的投票者——更确切地说，就是位于百分之五十位置上的那个选民。

在这里，一个完整的博弈包括两个阶段。在第一阶段中，候选人选择他们各

自在政治光谱中的位置。在第二个阶段，选民投票选出一位候选人。一般的两阶段博弈是允许之前所讨论的各种各样偏好的策略性不实陈述的。因此，对于我们的分析来说，为了避免在均衡中出现这样的行为，我们把候选人的数目减少到两个。在只有两位候选人的情况下，两阶段投票才会直接地对应于选民的偏好，而候选人在第一阶段的位置决策仍然保留了大博弈里真正有意思的那一部分。正是在那一阶段中，中位投票者定理限定了纳什均衡行为。

15.6.1 离散的政治光谱

让我们首先来考虑一场有着 9 000 万选民的选举，其中每个选民都在一种上面有五个点的政治光谱上有一个最偏好的位置（即政治立场）：极左（FL）、左（L）、中（C）、右（R）和极右（FR）。我们假设这些选民们围绕政治光谱的中心位置对称地分布着。他们的位置的**离散性分布**（the discrete distribution）由图 15—2 中的**柱状图**（histogram）或条形图表示。每个条形图的高度表示分布在那个位置上的选民的数量。在本例子中，我们假定，在 9 000 万选民中，4 000 万在左的位置上，2 000 万在极右的位置上，以及在极左、中和右等位置上的选民都各为 1 000 万。

在选举中，选民会投那些公开地将自己置于与他们自己在谱系中的位置较近的候选人的票。若两个候选人都与一群有着相同政治看法的选民等距离的话，则每位选民就通过抛硬币的方式来决定选择哪一个候选人。这一过程给予每个候选人该人群中一半的选票。

现在假定，一场总统竞选即将到来，候选人为克劳迪娅和多洛蕾丝，现在她们各自在竞选总统。[19]在图 15—2 所示的选民组成情况下，我们可以为两个候选

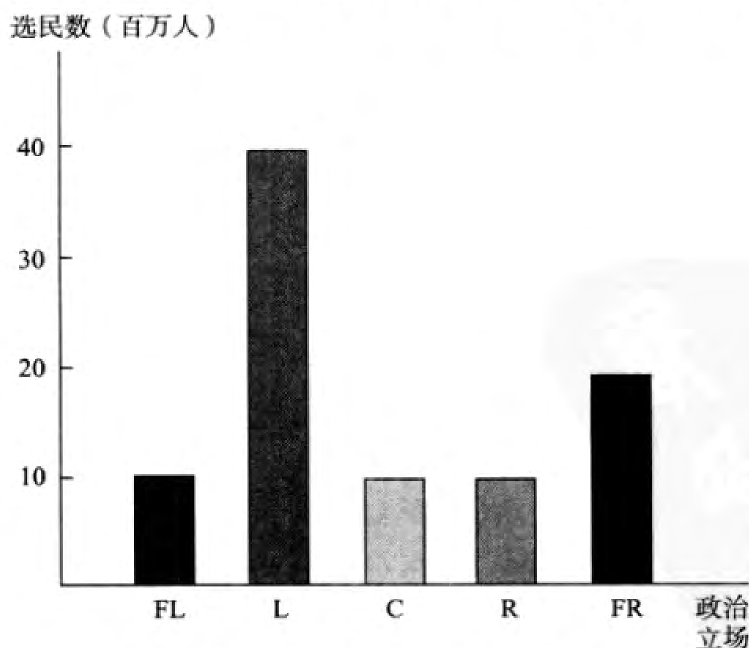


图 15—2　选民的离散分布

人构造一个支付表，以说明每一个候选人在所有的政治纲领选择的不同组合下各自预期能获得的票数。该 5×5 的表格如表 15—6 所示，以百万选票数为单位。候选人将会选择她们的最优位置策略来使她们所得到的选票数最大化（从而增加胜出的机会）。[20]

表 15—6　　　　候选人的政治立场博弈的支付表

		多洛蕾丝				
		FL	L	C	R	FR
克劳迪娅	FL	45，45	10，80	30，60	50，40	55，35
	L	80，10	45，45	50，40	55，35	60，30
	C	60，30	40，50	45，45	60，30	65，25
	R	40，50	35，55	30，60	45，45	70，20
	FR	35，55	30，60	25，65	20，70	45，45

下面对选票是如何配置的加以说明。当两个候选人都选择相同的位置时（沿着表中的左上角到右下角的对角线的五个单元格），每个候选人都将得到恰好一半的选票；因为所有选民与两个候选人都是等距离的，因此他们都通过抛硬币的方式来决定其选择，并且每个候选人都获得 4 500 万张选票。当两个候选人选择不同的位置时，位置较左的候选人会得到其位置上或其位置左边所有的选票，而位置较右的候选人会得到其位置上或其位置右边所有的选票。此外，每个候选人会获得相比于其对手更加靠近自己的中间位置的选票，而她们将平分与她俩等距的中间位置的选票。所以，如果克劳迪娅将自己定位于 L，而多洛蕾丝将自己定位在 FR，则克劳迪娅会在 L 上获得 4 000 万张选票，在 FL 上获得 1 000 万张选票，在 C 上获得 1 000 万张选票（因为 C 比较靠近 L 而不是接近 FR)。多洛蕾丝则会在 FR 上获得 2 000 万张选票，在 R 上获得 1 000 万张选票（因为 R 比较靠近 FR 而不是接近 L)，支付为（60，30)。类似的计算决定了表中其他单元格中的结果。

虽然表 15—6 看起来很大，但这个博弈却能很快地被求解。我们从现在已熟悉的二人博弈的占优策略或者劣策略的寻找开始。我们立即会看到，对于克劳迪娅来说，FL 劣于 L，并且 FR 劣于 R。而对于多洛蕾丝也有 FL 劣于 L 和 FR 劣于 R。剔去这些极端的策略后，对于每个候选人来说，R 劣于 C。当把这两个 R 剔除后，C 对于每一个候选人来说都是劣于 L 的。表中唯一剩下的单元格就是 (L，L)；这就是纳什均衡。

我们现在注意到在候选人的立场定位博弈的均衡的三项重要特性。第一，两个候选人在均衡时都位于相同的位置。这说明了**最小差别原则**（the principle of minimum differentiation)，它是所有的两人立场定位博弈中一般性结果，不管是总统候选人的政治纲领选择，还是街头小贩对热狗车位置的选择，或者电器制造商对产品特色的选择，都是如此。[21] 当投你票的人或者是向你购买货物的人可以被安排在某个指定的偏好光谱上时，你最好就是尽量做得与你的对手所做的完全

一样。这就解释了政治候选人的行为和商业行为趋于集中的事实。比如，它有助于你理解为何在繁忙的交通十字路口绝不会只有一家加油站，为什么即使每个牌子的车商都声明要推出“新款”车型，但所有牌子商家的四门箱式小客车（或迷你厢型车或运动休闲车）看起来都是一样的。

第二，可能也是更为关键性的，两个候选人都位于中位选民的位置。在我们的例子中，总共有 9 000 万选民，无论你从哪一端开始，中位选民都是 4 500 万。处于某个位置上的选民数可以任意地赋值，但是中位选民的位置是清楚的。在这里，中位选民位于政治光谱上的 L 位置。所以那就是两个候选人将他们自己设定的位置，这是由中位投票者定理所预测的结果。

第三，我们观察到，中位选民的位置并不一定要与光谱的几何中心相一致。如果选民的分布是对称的，这两个位置会一致；如果选民的分布偏向左边，则中位选民将位于几何中心的左边（如图 15—2 所示）；如果该分布偏向右边，那么中位选民将位于几何中心的右边。这有助于解释，比如，为什么马萨诸塞州的州政治候选人比在得克萨斯州或加利福尼亚州的处于类似位置的候选人更具自由倾向。

中位投票者定理可以用不同的方式来表达。一种版本将之简单地阐述为，中位选民的位置就是有两个候选人的竞选中的候选人均衡位置。另一种版本的表达则是，中位选民最为偏好的位置将会是孔多塞赢家；该位置将会在配对对决中击败每一个其他的位置。例如，如果 M 为中位位置，L 为 M 左边的任何位置，则 M 将会得到所有那些其最偏好位置位于 M 或者是在 M 右边的人们的选票，再加上部分在 M 左边但是距离 M 较近而距离 L 较远者的选票。因此 M 将获得 50% 以上的选票。两种版本都会产生相同的结果，因为在有两个候选人的选举中，两个试图赢得过半数选票的候选人都会选择孔多塞赢家位置。这些理解是完全相同的。另外，为保证该结果对于特定选民群体是成立的，该定理（不管是哪一种表达形式）要求每一选民的偏好是“合理的”，正如前面所要求的那样。在这里，所谓“合理的”意味着“单峰的”（single-peaked），正如在 15.3.1 节和图 15—1 中所描述的布莱克条件里的“单峰”概念一样。每位选民在政治光谱上都有唯一的最偏好的位置，而其效用（或支付）在任意一个方向上随着离此位置距离的增加而递减。[22] 在实际的美国总统竞选中，主要候选人都会向选民做出非常类似的承诺，这一倾向证实了中位投票者定理。

15.6.2 连续的政治光谱

中位投票者定理也可以在连续的政治立场分布下得到证明。**连续分布**（continuous distribution）假定有效地存在着无限多种政治立场，而不是只有 5 种、3 种或者任何有限数目的立场可供选择。这些政治立场与数轴上从 0 到 1 之间的位置相联系。[23]

选民依然和前面一样是沿着政治光谱分布的，但由于现在的分布是连续的而非离散的，所以我们使用选民**分布函数**（distribution function）而不是柱状图来

表示选民的位置。两种常见的分布——**均匀分布**（uniform distribution）与**正态分布**（normal distribution）——如图 15—3 所示。[24]每条曲线下的区域代表可获得的选票总数；在 0 到 1 区间中的任何给定点上，比如图 15—3（a）中的 x，计算出 0 到 x 之间分布函数曲线下面的面积，就知道直到此点为止的选票数量。很清楚，每一种分布的中位选民都位于光谱的中心位置，即位于 0.5 处。

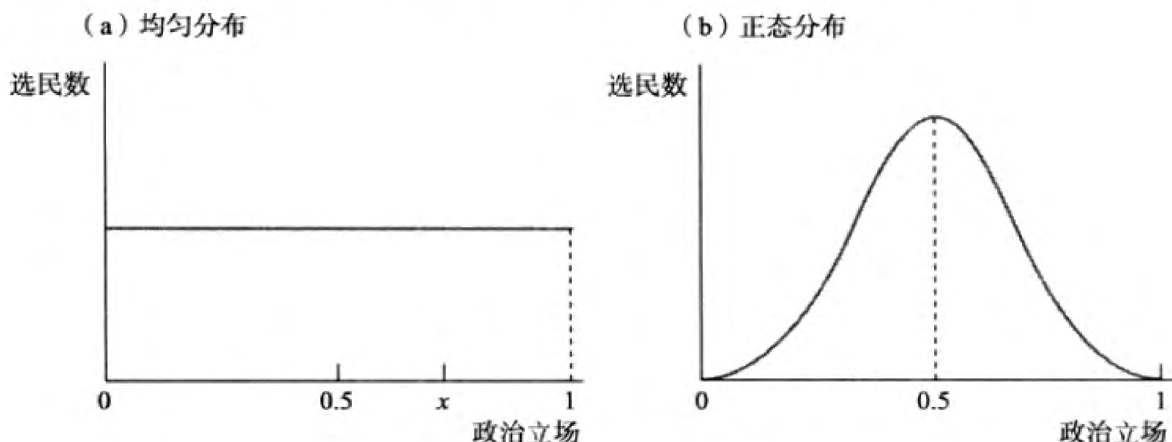


图 15—3　连续的选民分布

在连续光谱情形里，我们不可能为两个候选人构造一张支付表。这样的表格只能是有限维的，从而不能为参与人提供无限多可能的策略。然而，我们却可以用在 15.5.1 节中所讨论过的曾用于离散（有限）情形的同样的策略逻辑来求解该博弈。

考虑克劳迪娅和多洛蕾丝在思考所面临的可能的政治立场时的选择。每一个都知道她必须找到其纳什均衡策略——也就是对其对手的均衡策略的最佳反应。尽管可能的策略之完备集合是不可能加以描述的，但我们可以轻易地定义作为最佳反应的一个策略集合。

假设多洛蕾丝位于政治光谱上的一个随机的位置，比如图 15—3（a）中的 x，那么克劳迪娅就可以计算出她所有可能位置上选票将如何分配。若她选择 x 左边的位置，则她能够获得所有位于其左边的选票，以及位于她的位置与 x 之间的一半的选票。若她选择 x 右边的位置，则她能够得到所有位于其右边的选票，以及位于她的位置和 x 之间的一半的选票。最后，若她也选择位于 x 的位置，则她将与多洛蕾丝平分选票。给定多洛蕾丝的选择位于 x，则这三种可能性有效地归纳出了克劳迪娅所有的位置选择。

但是，刚才提出来的策略中的哪一个是克劳迪娅的“最优”反应呢？答案取决于 x 相对于中位投票者的位置。如果 x 在中位投票者的右边，则克劳迪娅知道其最优反应会最大化其所获得的选票的数量，她可以通过无限接近于 x 的左边来做到这一点。[25]此时，她有效地获得了从 0 到 x 的所有选票，而多洛蕾丝则得到所有从 x 到 1 的选票。如图 15—3（a）所示，当 x 位于中位投票者的右边时，则从 0 到 x 的分布曲线下面的面积所表示的选民数量据定义是大于从 x 到 1 分布曲线下的面积所表示的选民数量的；所以克劳迪娅会赢得选举的胜利。类似地，若

x位于中位投票者的左边，克劳迪娅的最佳反应将是定位在x右边非常近的位置上，并且获得从x到1的所有选票。当x刚好是在中位投票者的位置上时，克劳迪娅最好也选择在x。对于多洛蕾丝来说，其最优反应策略可以完全按照相同的方法来构造，并且给定其对手的位置，多洛蕾丝的最佳反应策略与已经描述的克劳迪娅的最优反应策略一样。在几何上，最佳反应曲线围绕在45°线周围并经过中位选民的位置，在该点上，这些最佳反应曲线恰好位于45°线上（当多洛蕾丝位于中位选民的位置时，克劳迪娅对多洛蕾丝位置的最佳反应就是定位于同一位置；相同的结论对于多洛蕾丝也是同样成立的）。

我们现在完成了对两个候选人最佳反应策略的描述。纳什均衡出现在最佳反应曲线的交点处，这种交点位于中位投票者的位置。通过以下方式，你可以很直观地得到这个结论。任意挑选一个候选人，将其置于任意一个开始位置，然后反复运用最佳反应策略，每个候选人最终会达到一个最佳位置。这个最佳位置即其应对其对手选择的位置的最优反应位置。如果多洛蕾丝考虑定位于图15—3（a）中的x处，那么克劳迪娅就会选择恰好在x左边的位置上，但是这时多洛蕾丝又会选择刚好在克劳迪娅左边的位置上，如此等等。只有当两个候选人都刚好位于分布的中间位置上时（无论分布是均匀的或者是正态的或其他某种类型）时，他们才会发现他们的决策对于彼此都是最优反应。我们又再一次看到，纳什均衡就是两个候选人都选择在中位选民的位置上。

对于真正的数学家来说，证明连续版本的中位投票者定理还需要更加复杂的数学方法。然而，就我们的目的来说，这里所给出来的讨论会使得你相信该定理在离散和连续情况下都是成立的。中位选民定理最重要的局限是，它只适用于只有一个议题的情形，或只适用于政治差异的一维光谱情形。如果存在着二维或者更多维时——例如，在社会议题上保守派与自由派的对抗与在经济议题上保守派与自由派的对抗并不一致时，人群就会分布于二维的“议题空间”中，而中位选民定理也就不再成立了。每一个个体投票者的偏好可以是单峰的，意思也就是说每个个人有其唯一的一个最偏好点，并且其支付值由此点沿着各个方向都是递减的，如同高度由山顶开始下降一样。但是，我们却无法在二维的情形里找出一个中位选民来，使得在中位投票者的位置的一边与另一边有同样数量的选民有其最偏好点。在二维的情形里，没有“边”的唯一概念，并且在两边的选民的数量是可以变化的，该变化取决于我们是如何定义“边”的。

15.7 总　　结

选举可以采用各种各样的不同投票程序来进行，这些不同的投票程序可以改变议题被考虑的顺序，或改变选票的统计方法。投票程序被分类为二元法、相对多数法或混合方法。二元法包括绝对多数规则和配对程序，如孔多塞方法和修正程序。相对多数规则和博尔达计数法和赞成投票这样的定位方法都属于相对多数的方法。而过半数决胜投票、即时决胜方法和比例代表投票法都是混

合法。

投票悖论（如孔多塞悖论、议程悖论与颠倒悖论）阐明了违反直觉的结果是如何产生的。悖论的产生或是由于加总偏好的困难，或是由于所考虑的议题列表的微小变更。另外一种矛盾性结果是在任何给定的选举中给定的选民偏好集可改变的情形下所产生的结果，它取决于所使用的投票程序。虽然阿罗不可能性定理表明没有一个规则能够同时满足所有的准则，但描述评估投票方法的一定的原则还是可行的。来自许多领域的研究者们都考虑过针对阿罗所给出的原则提出其他的替代方案。

选民们在选择投票程序的博弈中或在选举本身中有许多机会通过不实陈述其偏好进行策略性行为。偏好可以被策略性地加以不实陈述以获得投票者最喜欢的结果或者避免其最不喜欢的结果出现。在存在着不完美信息的情形，选民可以决定是否在其关于其他人的行为的信念以及他们对偏好分布的知识的基础上进行策略性的投票。

候选人在提出政治纲领时也可以表现出策略性行为。一个被称为中位投票者定理的一般结果表明，在仅有两个候选人的选举中，两个候选人都在中位选民的偏好位置上。当选民沿偏好光谱离散或连续地分布时，该结果都是成立的。

关键术语

议程悖论（agenda paradox）
修正程序（amendment procedure）
反相对多数方法（antiplurality method）
赞成投票（approval voting）
二元法（binary method）
布莱克条件（Black's condition ）
博尔达计数法（Borda count）
孔多塞方法（Condorcet method）
孔多塞悖论（Condorcet paradox）
孔多塞意义（Condorcet terms）
孔多塞赢家（Condorcet winner）
连续分布（continuous distribution）
柯普兰指数（Copeland index ）
离散分布（discrete distribution）
分布函数（distribution function）
吉巴德—萨特斯韦特定理（Gibbard-Satterthwaite theorem）
柱状图（histogram）
不可能性定理（impossibility theorem）
即时决胜（instant runoff）
不可传递顺序（intransitive ordering）
绝对多数规则（majority rule）
过半数决胜（majority runoff）
中位投票者定理（median voter theorem）
中位投票者（median voter）
混合方法（mixed method）
多阶段程序（multistage procedure）
正态分布（normal distribution）
配对投票（pairwise voting）
相对多数规则（plurality rule）
相对多数方法（plurative method）
定位方法（positional method）
最小差别原则（principle of minimum differentiation）
比例代表（proportional representation）
颠倒意义（reversal term）
颠倒悖论（reversal paradox）
稳健性（robustness）

轮（rounds）	分肥者（spoiler）
诚实投票（sincere voting）	不实的偏好策略性陈述（strategic misrepresentation of preferences）
单一可传递投票（single transferable vote）	策略性投票（strategic voting）
单峰偏好（single-peaked preferences）	可传递顺序（transitive ordering）
社会排序（social ranking）	均匀分布（uniform distribution）

习　题

1. 住在一个三人间宿舍里的三个室友 A，B 与 C 之间进行一次投票。他们想决定这学期要一起选修三门选修课程中的哪一门（每个人的专业都不相同，而每个人必须选几门专业课）。他们的选择有哲学、地质学和社会学，而他们对这三门课的偏好如下表所示。

A	B	C
哲学	社会学	地质学
地质学	哲学	社会学
社会学	地质学	哲学

他们已经决定要举行两轮投票，并以抽签方式来决定由谁来设定议程。假定由 A 来设定议程，而他想选择哲学课程。若他知道大家都会在每一轮中按真实意愿投票的话，那么他应该如何设定议程以获得他想要的结果？若他知道大家都会进行策略性投票的话，那么又应该采用什么样的议程呢？

2. 若由 B 负责设定议程而他又想要保证让社会学胜出，在这样的情况下重做习题 1。
3. 假定从选民 1 到选民 4 都被要求在一次博尔达计数法选举中对三位不同候选人 A，B 和 C 进行投票。他们的偏好如下表所示：

1	2	3	4
A	A	B	C
B	B	C	B
C	C	A	A

假定投票者会按照其真实意愿进行投票（没有策略性投票）。试找出一种博尔达计数规则使 A 获胜（在 A 获胜的情况下，第一、第二和第三偏好的候选人

分别被分配到多少分?)

4. 在候选人 B 获胜的情形下，请重做第 3 题，找出一个博尔达计数规则。

5. 考虑有 50 位居民出席在马萨诸塞州举行的一个小镇会议。他们必须在三种解决小镇垃圾问题的建议之间做出选择。建议 1 要求小镇把提供垃圾收集的服务作为其应该提供的服务之一；建议 2 要求小镇雇用私人垃圾收集者来提供垃圾收集的服务；建议 3 要求居民自己负责自家的垃圾收集。存在着三种类型的投票者。第一种类型偏好建议 1 胜于偏好建议 2，偏好建议 2 胜于偏好建议 3；这样的投票者有 20 位。第二种类型偏好建议 2 胜于偏好建议 3，且后者又胜于建议 1；这样的投票者有 15 位。第三种类型偏好建议 3 胜于偏好建议 1，且后者又胜建议 2；这样的投票者有 15 位。

   (a) 在相对多数投票规则下，何种建议会赢?

   (b) 假定投票的进行中采用了博尔达计数法，其中投票者按照其偏好的顺序将建议列在其选票上。选票上排在最前面（或顶端）的建议获得 3 分；列于第二位的获得 2 分；而列在最后的得 1 分。在这种情况下，如果没有策略性投票，每种建议分别可获得几分？哪种建议会赢呢?

   (c) 第二种与第三种类型的投票者可以运用什么样的策略，从而改变（b）中博尔达计数投票结果，将其变成两种类型投票者都偏好的结果？若他们运用此策略，每种建议分别会得到多少分？且哪种建议会赢呢?

6. 奥林匹克滑冰选手在其比赛中要完成两个项目，其一是短跑道比赛，另一个是长跑道比赛。在每一个项目中，滑冰选手们都被评分并且由 9 人裁判组进行排名，并且他们在排名中的位置被用于确定其最终得分。一个滑冰选手的排名取决于有多少裁判把其放在第一位（或第二位或第三位）；如果大多数裁判认为某个滑冰选手是最好的，那么该选手将排名第一，依此类推。在统计滑冰选手的最后得分时，短跑道项目权重为长跑道项目权重的一半。也就是说，最后得分＝0.5×短跑道中的排名＋长跑道中的排名。得到最低最终得分的滑冰选手将赢得金牌。在平局的情况下，在长跑道项目中被最多裁判评定为最佳的滑冰选手将拿到金牌。在 2002 年盐湖城（Salt Lake City）女子花样滑冰比赛中，在短跑道项目结束之后，米歇尔·凯文（Michelle Kwen）排在第一位，紧随其后的是伊里娜·斯鲁茨卡娅（Irina Slutskaya）、莎莎·科恩（Sasha Cohen），以及萨拉·休斯（Sarah Hughes），她们分别处于第二、第三和第四位。在长跑道项目中，裁判给这四个选手的评分如下表所示。

   (a) 在奥林匹克竞赛中，斯鲁茨卡娅是顶级滑冰选手中的最末一位。使用从裁判卡片中得到的信息在斯鲁茨卡娅滑冰之前决定裁判们在长跑道中给凯文、科恩和休斯的排名。然后，使用从短跑道中已经给定的排名，结合你已经计算出来的长跑道排名，试决定在斯鲁茨卡娅滑冰之前，其他三个选手的最后得分以及排名顺序。（注意凯文在短跑道中的排名是 1，故她在短跑道后的部分得分是 0.5）。

   (b) 给定你对（a）的答案，如果裁判们在长跑道中把斯鲁茨卡娅排在其他三名选手之前，那么比赛的最终结果会是怎样的呢?

   (c) 在斯鲁茨卡娅滑冰之后，使用裁判的卡片来确定所有四位选手最终的实

际得分。每块奖牌分别由谁获得？

(d) 在阿罗所给出来的原则中，什么样的重要原则在奥林匹克花样滑冰计分规则中被违反？试给予说明。

		裁判编号								
		1	2	3	4	5	6	7	8	9
凯文	分数	11.3	11.5	11.7	11.5	11.4	11.5	11.4	11.5	11.4
	名次	2	3	2	2	2	3	3	2	3
斯鲁茨卡娅	分数	11.3	11.7	11.8	11.6	11.4	11.7	11.5	11.4	11.5
	名次	3	1	1	1	4	1	2	3	2
科恩	分数	11.0	11.6	11.5	11.4	11.4	11.4	11.3	11.3	11.3
	名次	4	2	4	3	3	4	4	4	4
休斯	分数	11.4	11.5	11.6	11.4	11.6	11.6	11.3	11.6	11.6
	名次	1	4	3	4	1	2	1	1	1

7. 在古巴导弹危机期间，关于为约翰·肯尼迪总统提供咨询的国家安全委员会执行委员会（ExComn）内部所出现的严重意见分歧，我们在这里归纳如下。存在着三种选择：温和的反应（封锁）、中度的反应（有限度的空袭）以及强硬的反应（大规模的空袭或入侵）。在国家安全委员会执行委员会中有着三个集团。鸽派认为最好的选择是选择温和的反应，其次才是中度的反应，排在最后的是强硬的反应。鹰派最偏好于中度的反应，其次是强硬的反应，排在最后的是温和的反应。军方最偏好强硬的反应，但是他们觉得“有限度的空袭”所带来的危险太大，以至于他们宁可不采取任何的军事行动，也不愿采取“有限度的空袭”（见 *The Kennedy Tapes*：*Inside the White House During the Cuban Missile Crisis*，ed. Ernest R. May and Philip D. Zelikow，Cambridge，MA：Harvard University Press，1997）。换句话说，他们将温和的反应置于第二位，将中度的反应置于最后。每个集团大约占了国家安全委员会执行委员会全部成员的三分之一，因此，任何两个集团都可以形成多数。

(a) 如果在国家安全委员会执行委员会中由绝对多数规则来决定做何选择，并且其中的成员都诚实地投票，那么什么样的选择——如果有的话——会胜出呢？

(b) 如果成员们进行策略性投票的话，会出现什么样的结果？如果有一个集团有设定议程的权力，又会出现什么样的结果？（阅读 15.2.2 节和 15.4.2 节中的分析之后，对你在这两个案例里的讨论建模。）

8. 约翰·鲍罗斯（John Paulos）在其著作《读报的数学家》（*A Mathematician Reads Newspaper*）中，基于 1992 年民主党总统初选会议，做出了如下讽刺。有五位候选人欲代表民主党角逐总统大选：杰里·布朗（Jerry Brown）、比尔·克林顿（Bill Clinton）、汤姆·哈金（Tom Harkin）、鲍勃·克里（Bob Kerrey）

与保罗·桑加斯（Paul Tsongas）。有55位投票者，各自对所有候选人有着不同的偏好顺序。存在着6种不同的偏好顺序，我们以I到VI来表示。下表给出了偏好顺序（从最好的1到最差的5）以及每一种顺序的投票者数目；候选人以姓氏的第一个字母来表示。[26]

		投票方案及投票者数量					
		Ⅰ18	Ⅱ12	Ⅲ10	Ⅳ9	Ⅴ4	Ⅵ2
排序	1	T	C	B	K	H	H
	2	K	H	C	B	C	B
	3	H	K	H	H	K	K
	4	B	B	K	C	B	C
	5	C	T	T	T	T	T

(a) 首先假定所有投票者都诚实地依其偏好进行投票，考虑几种不同选举规则下的结果。试给出下列每一种结果：（Ⅰ）在相对多数方法下（最为偏好的那一位），桑加斯赢。（Ⅱ）在对决方法下（最偏好的前两位进入第二轮），克林顿赢。（Ⅲ）在剔除方法下（剔除每一轮中偏好序位最低者，剩下的进入下一轮），布朗赢。（Ⅳ）在博尔达计数法下（最偏好者可获得5分，第二获4分，依此类推；而获最高分的候选人赢），克里赢。（Ⅴ）在孔多塞方法下（配对比较），哈金赢。

(b) 假设你就是布朗、克里或哈金的支持者。在相对多数方法下，你会得到最糟糕的结果。你是否能从策略性投票中获得好处？如果是这样的话，应该怎样做？

(c) 在每一种其他方法下，也存在进行策略性投票的机会吗？如果是的话，试说明谁能从策略性投票中获利以及他们如何去实施？

9. 回忆在15.4.3节中所讨论的三人议会考虑三个可选择的福利政策的例子。在那里，三个议员（“左”、“中”和“右”）在第一轮投票中考虑政策A和D，其胜出者则会在第二轮投票中与G对决，但是没有人确切地知道每一个可能的偏好集有多少议员；可能的偏好顺序在表15—1中已经给了出来。每一个议员都知道他自己的类型并且知道观察到其他议员的类型（“左”、“中”、“右”）的概率 p_L、p_C 和 p_R（$p_L+p_C+p_R=1$）。在第一轮选举中，中类型议员的行为在这种情况下是唯一不为人所知的，并且它取决于各种偏好类型出现的可能性。这里假定中类型议员相信（与正文中所考虑的情形相反）其他的中类型议员将会进行策略性投票；进一步假定，中类型议员的支付如15.4.3节所述：如果A赢就为1，如果G赢就为0，如果D赢就为 $0<u<1$。

(a) 在其他两张选票的什么样的配置下，中类型议员的第一轮投票会影响选举结果？给定他对其他中类型议员行为的假定，他会如何识别出首轮选票的来源？

(b) 根据15.4.3节中的分析，试确定中类型议员诚实投票时的期望支付。将

其与他进行策略性投票时的期望支付进行比较。在什么条件下，中类型议员会进行策略性投票？

10. 一次选举中有三个候选人且该选举采用的是相对多数规则。存在着大量的投票者，从左到右分布在一种意识形态光谱上。用一条从0（左）到1（右）的水平直线来表示这一分布。投票者沿此光谱而均匀地分布着；所以在该直线上的任意区间的投票者数量就与该区间的长度成正比。所以 有1/3的投票者分布在0到1/3的线段上，1/4的投票者分布在1/2到3/4的线段上，依此类推。每个投票者都会投票给其宣称的立场最接近于投票者自己立场的那个候选人。候选人没有自己的意识形态追求，他们可以采取直线上的任何立场，每一个都仅仅寻求最大化其选票份额。

(a) 假定你是三个候选人中的某一个。其他两个候选人中，最左边者位于 x，最右边者位于 $(1-y)$，其中 $x+y<1$（所以最右边者离1有 y 的距离）。试证明在给定条件下你的最优反应就是采取下列立场：

(ⅰ) 若 $x>y$ 且 $3x+y>1$，稍微靠近 x 的左边；

(ⅱ) 若 $y>x$ 且 $x+3y>1$，稍微靠近 $(1-y)$ 的右边；

(ⅲ) 若 $3x+y<1$ 且 $x+3y<1$，刚好位于其他两位候选人的中间。

(b) 在以 x 和 y 为坐标轴的图上，标示出表示你每个最优反应规则［(a) 中的 (ⅰ) 到 (ⅲ)］的区域（x 与 y 值的组合）。

(c) 通过你的分析，关于其中三个候选人每人都选择立场的博弈的纳什均衡，你可以得出什么样的结论呢？

【注释】

[1] 关于此主题的经典教科书，是曾被用于在政治科学中普及博弈论工具的由威廉·瑞克尔（William Riker）所著的《自由主义与平民主义》(*Liberalism Against Populism*, San Francisco：W. H. Freeman，1982)。一般性的综述见《投票经济学》（"Economics of Voting"，*Journal of Economic Perspectives*，Vol. 9，no. 1，Winter 1995)。一个重要的早期贡献是密契尔·都默特的《投票程序》(Michael Dummett，*Voting Procedures*，Oxford：Clarendon Press，1984)。唐纳·沙利（Donald Saari）的《混乱的选择》(*Chaotic Elections*，Providence，RI：American Mathematical Society，2000)，提出了我们在本章中随后要用到的一些新思路。

[2] 注意，这种指数或得分一定要有精确的机制去处理平局。世界杯足球赛采用了一种低估平局价值的规则，以鼓励更具进攻性的比赛。见 Barry Nalebuff and Jonathan Levin，"An Introduction to Vote Counting Schemes，" *Journal of Economic Perspectives*，vol. 9，no 1，Winter 1995，pp. 3-26。

[3] 不像其他许多需回溯过去好几个世纪历史的方法，赞成投票方法是由研究生罗伯特·韦伯（Robert Weber）在1971年设计和命名的。韦伯现在是西北大学管理经济学与决策科学方面的专家，擅长博弈论。

[4] 它是如此有名，以至使得人们都知道经济学家们称其为"投票悖论"（the voting paradox)。政治科学家似乎知道得更多，因为他们更为经常地使用其正式名称。正如我们看到的那样，存在大量的投票悖论，不仅仅限于用孔多塞的名字命名的这一个。

[5] 参见 Peter Ordeshook，*Game Theory and Political Theory*，Cambridge，UK：

Cambridge University Press, 1986, p. 58。

[6] 该定理的完整表达，通常被称为"阿罗一般可能性定理"（Arrow's General Possibility Theorem），可参见 Kenneth Arrow, *Social Choice and Individual Valves*, 2nd ed., New York: Wiley, 1963。

[7] 参见 Walter Nicholson 在其《微观经济理论》（*Microeconomic Theory*, 7th ed., New York: Dryden Press, 1998 pp. 764 - 766）中对阿罗不可能性定理更为详细的论述，适合中级经济学水平的学生阅读。

[8] 参见 Duncan Black, "On the Rationale of Group Decision-Making," *Journal of Political Economy*, vol. 56, no. 1, 1948, pp. 23 - 34。

[9] 参见 Partha Dasgupta and Eric Maskin, "On the Robustness of Majority Rule,", working paper, April 2000。

[10] 关于萨里在阿罗不可能性定理上的工作的更多具体信息，参见 D. Saar "Mathematical Structure of Voting Paradoxes I: Pairwise Vote," *Economic Theory*, vol. 15, 2000, pp. 1 - 53。关于这个结果和博尔达计数法的稳健性的信息可参见 D. saari, *Chaotic Elections*, Providence, RI: American Mathematical Society, 2000。

[11] 请注意，赞成投票的方式就不会面临与此相同的问题。

[12] "Ross Returns," *Newsweek*, Special Election Recap Issue, November 18, 1996, p. 104.

[13] 在关于投票的文献中，有一个一般性的结果，即选民们在面对成对的候选人时，将在投票的最后一轮中总是进行不带欺骗性的真实的投票。

[14] 对于这种情形的更为完整的分析可参见 Riker, *Liberalism Against Populism*, pp. 152 - 157。

[15] 这种结果参见 P. Ordeshook and T. Palfrey, "Agendas, Strategic Voting, and signaling with Incomplete Information," *American Journal of Political Science*, vol. 32, no. 2, 1988, pp. 441 - 466。这个例子的结构基于 Ordeshook 与 Palfrey 的分析。

[16] 仅仅当其他所有选民都平均地在 A 和 D 之间分配其票数时，中类型才能够影响选举的结果。因此，在第一轮中一定要恰好有 $(n-1)/2$ 个右类型选民选择 D，并且有 $(n-1)/2$ 个其他选民选择了 A。如果那些选择 A 的选民为左类型，则 A 就不能够赢得第二轮选举而中类型就会得到零支付。为使中类型得到支付 1，就必须要求所有支持 A 的其他选民为中类型。这种情况发生的概率为 $[p_C/(p_L+p_C)]^{(n-1)/2}$。此时中类型按照其真实意愿进行投票时所得到的期望支付就是如这里所述的式子。见 Ordeshook and Palfrey, p. 455。

[17] 有关该结果的理论细节，见 A. Gibbard, "Manipulation of Voting Schemes: A General Result." *Econometrica*, vol, 41., no. 4, 1973, pp. 587 - 601，以及 M. A. Satterthwaite, "Strategy-Proofness and Arrow's Conditions," *Journal of Economic Theory*, vol. 10, 1975, pp. 187 - 217。该定理使用了他们两人的名字是由于两人都各自独立地证明了这个结果。

[18] H. Nurmi 的分类可以在其所著的 *Comparing Voting Systems* (Norwell, MA: D. Reidel, 1987) 一书中找到。

[19] 我们假定的候选人与美国过去实际上的或是将来可能的候选人之间的相似并不意味着暗含了对她们的表现与纳什均衡的关系的一种分析或预测。而我们的选民分布也不意味着是对美国选民偏好的典型刻画。

[20] 为了简化分析，我们忽略了由选举团所造成的复杂性，并假设只考虑大众投票。

［21］经济学家们在关于空间区位竞争的 Hotelling 模型中推导了这一结论。见 Harold Hotelling，"Stability in Competition，" *Economic Journal*，vol. 39，no. 1，1929，pp. 41－57。

［22］然而，选民沿着政治光谱的理想点分布不一定是"单峰"的，如图 15—2 中的柱状图就不是单峰的，它在 L 和 FR 处共有两个峰值。

［23］这种构造分布与第 12 章和第 13 章中分析由大量个体成员所组成的人群时所采用的构造分布是相同的。

［24］我们不去深入探究分布理论的潜在机制，或用于计算位于连续的政治光谱上任何特定位置的左边或右边的投票人的精确比例所需的积分方法。这里我们仅给出足够的信息让你相信，在连续的情形里，中位选民定理仍然是成立的。

［25］在连续的情形，这样一种位置，无限小地从 x 处移到左边是可行的，而在我们离散的例子中，候选人必须刚好处于相同的位置。

［26］John Allen Paulos，*A Mathematician Reads the Newspaper*，New York：Basic Books，1995，pp. 104－106.

第 16 章　投标策略与拍卖设计

本章我们讨论与日常生活息息相关的主题——投标策略与拍卖设计。你或许会争辩说你未曾参加过任何佳士得或苏富比的正式拍卖会，又或者伊利诺伊州的牲畜拍卖会，但实际上你参加了比你想象中更多的拍卖会。又比如，随着最近网络拍卖站点的流行，数百万人会定期参加拍卖购物。而你所参与的其他交易也许就是拍卖的一种类型。

拍卖需要将一件物品以某一价格（投标价）由卖方转移至买方（投标者）。在这种简单形式下，许多市场交易是与拍卖类似的。例如，波士顿的 Filene's Basement 商店就应用巧妙的价格策略，让顾客一直回来购买更多的商品：它每周都持续地对还在货架上的商品降价，降价会一直持续到此商品被买走；当价格降得太多时，它会干脆将商品捐献给慈善机构。顾客喜欢这样的方式。然而，很少有人真正知道，其实他们已参与了所谓的降价拍卖，或荷兰式拍卖。本章将会对这种类型的拍卖做详细说明。

即使你没有亲自参加拍卖，拍卖还是会深深地影响你的生活。1994 年以来，美国联邦通讯委员会（FCC）进行了不少于 40 场拍卖，拍卖了大部分的电磁波频段。你在之后数十年所观看的电视、所使用的移动电话都会受到这类过程及其结果的影响。

截至 2003 年 6 月，这类拍卖为政府带来了 410 亿美元的收入，而且 2003 年夏秋之间还安排有 4 场拍卖会。由于这笔收入为联邦预算带来很大的贡献，其对宏观经济总量已产生重大影响，比如在利率方面。国际经济不仅受到美国电

磁波频段拍卖的影响，而且也受到至少六个欧洲国家以及澳大利亚与新西兰类似拍卖的影响。明确拍卖如何进行，有助于你了解这类重要的事件及其可能产生的影响。

从策略的视角看，拍卖具有若干个有趣的特征，最关键的是买卖双方之间以及投标者之间存在信息不对称，这样，信号传递和甄别机制就成为买卖双方策略中的重要部分。此外，买卖双方的最优策略将依赖于他们的风险规避程度。我们也将发现在某些特殊情况下，对于若干类型的拍卖而言，卖方以及投标赢家的期望支付是相同的。

本章将探讨你可能会碰见的各种类型的拍卖，以及在相应情况下买卖双方可能会使用的策略。虽然正式拍卖理论需要用到高等微积分来推导结果，但是在这里我们会跳过大部分困难的数学计算，而利用直觉说明最佳行为与策略的选择。[1]我们将以网上拍卖为例阐释拍卖设计与投标理论是如何影响现代商业的，并提供关于对此类拍卖的详细讨论。

16.1 拍卖的类型

不同类型的拍卖在喊价的方式及投标赢家最终付款的标准上各不相同。此外，拍卖还可以按照买方评估拍卖品价值的方式进行分类。这里我们区分各种拍卖类型，并描述其特性与机制。

四种主要拍卖类型可以归结为两大类。第一大类称为**公开喊价**（open outcry）。在这种拍卖中，投标方在公开场合喊价，所有投标者都能够观察到投标价。这种类型的拍卖也许最符合流行的对拍卖的理解。在人们看来，拍卖总是与狂热的投标者与拍卖者联系在一起的；但是，公开喊价又可以分为两种方式，只有其中之一才可能出现“狂热的”投标。

升价拍卖（ascending）或**英式拍卖**（English）是公开喊价拍卖中最符合一般人对拍卖的印象的一种拍卖方式，是在英式拍卖会场中（如佳士得或苏富比）举行的标准式拍卖的另称。依惯例，英式拍卖会场中会有一位拍卖者，他会从一个低价开始进行拍卖，接着会喊出越来越高的价格。在继续喊出下一个价格之前，拍卖者会等待是否有投标者要投标。当没有投标者继续出价时，此拍卖品就由出价最后最高的投标者获得。这样，任何投标者都可参与英式拍卖，虽然只有出价最高的投标者能获得出售的物品。投标过程不必如实公开喊价，因为点头或手腕轻敲一下都可被视为拍卖中的叫价行为。现在，在互联网上，也有大量网站以类似英式拍卖的方式拍卖你能想象到的几乎任何实际商品。

另一种公开喊价拍卖是**荷兰式拍卖**（Dutch）或**降价拍卖**（descending）。荷兰式拍卖因郁金香等花卉在荷兰拍卖采用这种方式而得名。它与英式拍卖进行的方向相反，拍卖者以一个非常高的价格开始，然后喊出越来越低的价格，一直到有投标者接受该价格并进行投标，从而获得拍卖品。出于拍卖速度的需要，荷兰

式花卉拍卖，与其他农产品或易腐蚀商品（如悉尼每天鱼市场拍卖）拍卖一样，会使用“时钟”对每个投标价进行倒数计时，一直到有人“停住时钟”并取走商品为止。在许多情况下，拍卖时钟会泄露出相当多的有关当前出售的商品及降价的信息。不同于英式拍卖，荷兰式拍卖过程中没有狂热的喊价，原因在于只有一个投标者会最终“停住时钟”取走物品。

第二大类拍卖叫**密封投标**（sealed bid），投标秘密进行，投标者之间不能观察到其他人的投标价，在许多情况下，只宣布赢家的投标价。这种密封拍卖，如同荷兰式拍卖一样，投标者只有一次投标机会。（技术上，你可以多次叫价，但只有最高价才会与拍卖结果有关。）密封式拍卖不需要拍卖者，只需要一个监督员公布投标价并确定谁是赢家。

在密封投标拍卖下，决定中标者支付的价格的方法有两种：在**第一高价**（first-price）密封投标拍卖下，出价最高的投标者赢得物品并按其投标价支付；在第二高价密封投标拍卖中，出价最高的投标者赢得拍卖品，但是仅需支付第二高的投标价。**第二高价**（second-price）密封投标拍卖的安排可引出诚实的投标，这将在16.4节中讨论。这样的拍卖称为**维克里拍卖**（Vickrey auctions）。维克里是第一个发现这种特性的经济学家。我们也将看到上述密封投标在拍卖策略与期望支付上分别相似于某一种公开叫价拍卖：第一高价密封投标拍卖与荷兰式拍卖相似，而第二高价密封投标拍卖却与英式拍卖相似。

最后，投标者可以利用两种方法来评估拍卖品的价值。在**一般价值**（common-value）或**客观价值**（objective-value）拍卖中，拍卖品对于所有投标者而言，其价值是相同的，但是每个人都只知道一个不准确的估计值。投标者对其可能的分布有所了解，但每个投标者都必须在投标之前做出自己的评估。例如，一块蕴藏固定石油量的油气田对于所有公司而言，收益应该是相同的，但是每家公司对该油田所蕴藏的石油量却有各自专业的评估。同样地，每位债券交易人对未来利率也只有自己的估计值。在这种情况下，信号传递与甄别机制就显得举足轻重；每位投标者都清楚，其他投标者会有一些关于拍卖品价值的信息（即使是不完全的），而他应该从对手的行为中推测其拥有的信息内容。另外，他还需意识到他的行动也会把自己的私人信息传递给其他投标者。

私人价值（private-value）或**主观价值**（subjective-value）拍卖则属于第二种估价类型。在这种情况下，针对同一件拍卖品每位投标者会给出不同评价。例如，对于某些投标者而言，黛安娜的一件礼服或贾桂琳用过的一件首饰具有情感上的价值。投标者只知道该物品在自己心中的价值，但并不知其在别人心中的价值，同样地，卖方也无法知道投标者对该物品的估价，买卖双方可能彼此推测对方的估价，再利用信号传递与甄别机制修改他们的出价。信息的问题，不仅会影响到私人价值拍卖中买方的投标策略，也会影响到卖方的策略——如何设计拍卖形式以辨别最高估价并取得好的价格。

另外，也可以构造其他的拍卖方案。例如，你可以将拍卖设计成由出价最高的投标者获得物品，但是由所有投标者支付投标价，我们将于16.5节中讨论拍卖设计的程序。我们无法讨论所有组合，因此，我们会分析一些比较普遍的拍卖方案，并应用实例讨论重要的策略概念。

16.2 赢家的诅咒

一般价值的拍卖会引起一项常被忽略的标准结果。请回忆之前曾提到的一类拍卖，其中拍卖品对于所有投标者而言价值完全相同且固定，然而每位投标者只能估计其价值。**赢家的诅咒**（winner's curse）就是对投标者的警告，因为他们一旦赢得了拍卖品，就很可能付出了大于其价值的价格。

假设你是一个竞标塔吉可公司的外部并购者。你的专家已经对塔吉可公司进行了研究并对其价值做出了估计，在目前的经营状况下，该公司的价值界于 0 至 100 亿美元之间，所有在此范围内的值的可能性都是相等的。目前经营者知道其真正价值，他们却不会告诉你。你相信，无论塔吉可公司的价值在目前经营状况下为多少，在你的经营下其价值将会提升 50%，那么你该出价多少来投标呢？

你或许会想，塔吉可公司在目前的经营状况下，其平均价值为 50 亿美元，而在你的经营下，其平均价值将会达到 75 亿美元。这样，那么界于 50 亿至 75 亿美元之间的投标价应该是有利可图的；然而，此投标策略却没有考虑目前经营者对你投标价的反应，若塔吉可公司的实际价值大于你的投标价，则该公司目前所有者是不会接受你的投标价的。只有在其实际价值接近于你投标价范围的下限时，你才可能中标。

假如你的投标价为 b，在目前经营状况下，只有在塔吉可公司的价值界于 0 至 b 之间时，你的投标价才会被接受而你才能接手该公司的经营权；平均而言，若目前所有者接受你的投标价，你就可以估计此公司目前的价值是 $b/2$。在你的经营下，其平均价值将会比目前提升 50%，即 $(1.5)(b/2)=0.75b$。由于此价值小于 b，因此只有在不值得获胜的时候，你才会赢得其经营权！许多外部并购者发现这个事实时，都已经为时已晚。

外部并购者经常会与目标公司进行一对一的谈判，这很类似于只有一人在进行投标。其实，并非只有外部并购者会被赢家的诅咒所影响，当你与其他投标者在一般价值的拍卖中竞争，而各自对拍卖品做出独立估价时，也会产生相同的问题。

以一片油气田（或蕴藏油气的海域）的勘探权租约为例，在这场拍卖中[2]，只有在对手对油田的估计价值比你低时，你才会赢。你必须认清这点并从中得到教训。

假设没有任何投标者知道油气田的真正价值为 10 亿美元（在本例中卖方也可能不知道油田的真正价值。）假定有 10 家石油公司准备参与投标。每家公司的专家对此油田的估计误差都为 ±1 亿美元，所有在此范围内的数字都是等可能的。若对 10 家公司的估计值求平均值，则此值比起任何其他单一估计值都会更准确。然而当每位投标者只专注于自身的估计值时，大部分的估计值都会有偏差。这时的估计值，平均而言为 10.8 亿美元，非常接近估计值的上限。[3]因此，获胜的公司可能就多付了很多，投标公司只有认清问题，并将其投标价向下调整

到抵消误差处，才能解决这个问题。正确的计算必须要能确保你即使将投标价向下调整也不会失去这场拍卖，但是，这很困难，因为你必须意识到所有其他投标者也正在做相同的调整。

我们不会在一般价值的拍卖中运用高等数学来寻找最佳的投标策略，但是，我们可以提供给你一般性的建议。若你对某拍卖品进行投标，则请注意下面两个问题是完全不同的："在投标以前，根据我所掌握的信息，我会以10.8亿美元购买这块油田吗?""在投标以前，根据我所掌握的信息，且已知只要没有其他人愿意投标10.8亿美元，我就能够购买它，我还是愿意以10.8亿美元购买这块油田吗?"[4]即使是在密封拍卖中，第二个问题还是显示了正确的策略性思考，因为只有其他人的投标价都比你低——只有其他人对拍卖品的估价比你的估价低时，你才能获胜。

若你的投标行为没有考虑到赢家的诅咒，你应该想到你将会损失一些钱。之前对投标塔吉可公司所做的计算就说明了赢家的诅咒。实际上，这种危险真的会发生吗？里查德·泰勒（Richard Thaler）以许多事实证明了这种危险的确存在。[5]

我们可以利用拍卖一整罐硬币来进行赢家的诅咒的简易实验。价值是客观的，但是每位投标者要各自对两个问题形成主观上的估计：罐子里有多少硬币？总价值为多少？这是个一般价值拍卖的例子。许多老师都会引导学生做类似的实验，而都发现有出价过高的情况发生。在一个相似而相关的例子里，一群MBA学生被要求对一家虚拟公司进行投标，而不是一罐硬币。在每一个回合后，学生们都会得到关于公司的真正价值的反馈信息。博弈会一直重复进行。而69位学生里只有5位学生会每次都出价低一点，因此，愈到后面的回合，平均的投标价格愈高。

对实际现象的观察更证实了这些发现。有证据证明油气田勘探权的拍卖得主在租约上其实是有损失的。我们发现，比较转会到新球队的棒球选手与继续和球队续约的棒球选手，前者往往会得到超额的支付。

"考虑到赢家的诅咒，投标价要向下调整多少才合理?"回答这个问题超出了本书的范围；16.9节所引用的文章则提供了所需的数学分析，供你参考。我们在此仅指出问题的所在，并强调其重要性。特别是当你的支付意愿取决于你对交易获利能力的预期，或者是你对拍卖品再次售出时的价值的估计时，你就需要谨慎考虑。

本节分析显示了博弈理论预测作用的重要性。通过观察和实验的证明，我们知道有很多人陷入赢家的诅咒之中，因而损失了大笔金钱。学习基本的博弈理论有助于认识赢家的诅咒，并避免损失。

16.3 投标策略

我们现在转向私人价值拍卖，并讨论最佳的投标策略。假设你对1952年酿

制的法国葡萄酒特别感兴趣，请想一下你可以利用那些可能的拍卖程序来销售葡萄酒。

假设你正在参加一场标准的英式拍卖。在你知道你本身对拍卖品的估价为 V 的情况下，你的最佳投标策略就很明确了。你可以从投标过程的任何一步开始。若上一个投标价是由对手所提出的 r，而 r 刚好为 V 或者比 V 高，则你绝对不会想提出更高的投标价；因此你也不会关心任何进一步的投标状况。只有在上一个投标价仍然比 V 低时，你才会继续进行投标。在这个例子中，你可以增加一分钱（或者是拍卖会允许的最小增加额），以 r 再加上一分钱的价格进行投标。若拍卖在此价格结束，你就能以投标价 r（或者是实际价格 r）购得葡萄酒，而你能够获得 $V-r$ 的有效利润。若投标继续进行，你可以重复此步骤，一直在 r 基础上提高投标价。在此类拍卖中，出价最高者会以第二高价者的估计价格获得此葡萄酒，最终价格接近第二高价的程度取决于拍卖规则中所规定的最小增加额。

现在假设葡萄酒拍卖为第一高价密封投标拍卖，而你是一个估价很高的投标者。你需要决定是以 V 还是以别的价格进行投标。你是否应该以对拍卖品全部估价作为投标价格？

请记住，在这种拍卖中，中标者需要支付其投标价格。因此在此例中，你不应该以 V 来投标。这样的投标价格会让你无法获得任何利润，所以你应该降低投标价。但是，若你以略低于 V 的价格投标，而你的对手却以低于 V 但高于你的投标价的价格投标，那么你会有投标不中的风险。然而，只要你的投标价不是太低以致你无法中标，还是有很大可能获得利润的。你的最佳投标策略就是逐渐降低你的投标价。描述这一策略需要用到微积分，但是直观上理解这一结果也很简单。对于你而言，从 V 开始递减你的投标价有利有弊：若你获得葡萄酒，它就会增加你的利润，但是它同时也会降低你中标的机会，从而减少你获得葡萄酒的可能性。当递减到最后的投标价恰好平衡上述两种作用时，该投标价就是最理想的。

若以荷兰式拍卖的方式进行，结果会如何呢？在这种情况下，你的投标策略会类似于第一高价密封投标拍卖时的情形。请试想你的投标概率。当拍卖者喊出的价格比 V 高时，你会选择不再继续投标。当没有人以低于 V 的价格进行投标时，那么你也不会参与投标。但是，如果在密封投标拍卖中，你会有两种选择。你可以现在投标，而得不到任何利润，或是等价格低一点时再投标。再等待久一点，可以增加你从投标品得到的利润，但同时也会增加失去葡萄酒的风险。因此，你的利益在这里也跟着递减，而确切的递减值取决于前一段所描述的成本—收益分析。

最后，讨论采用第二高价密封投标拍卖时的情形。这时，关于逐渐降低投标价的成本—收益分析不同于前面三种拍卖。这个结果源于如下事实，从逐渐降低投标价中获得的利润为零，依靠递减投标价你不会获得利润，因为你的利润取决于第二高价，而不是你自己的叫价。鉴于第二高价密封投标拍卖所具有的趣味性，我们将在下一节分析其相关的投标策略。

16.4 维克里说真话规则

我们由上一节得知，私人价值英式拍卖与第一高价密封投标拍卖的投标策略是不同的。在英式拍卖中，出价最高的投标者能够以第二高的投标价格获得拍卖品。而在密封投标拍卖中，出价最高的投标者需要支付其投标价，而不管其与第二高价之间的价差是多少；在密封投标拍卖中，策略性的投标者明白这个事实，并企图利用递减其投标价的方式来获取利润。当然，在其他条件不变的情况下，卖方会偏好未经递减的投标价格，让我们看看卖方会如何诱导投标者在投标中泄露其真实估价。

以密封投标方式拍卖一项私人价值物品时，维克里（Willian Vickrey）表示，若卖方放弃第一高价密封投标方式而使用修改后的密封投标方式进行拍卖，那么将会有更多投标者泄露其真实估价；他建议修改密封投标拍卖的规则，使其更相似于公开喊价拍卖。[6]也就是说，出价最高的投标者以次高的投标价获得拍卖品，即第二高价密封投标拍卖。维克里认为若按这些规则进行拍卖，每位投标者的占优投标策略都是以真实估价进行投标。因此我们将其戏称为**维克里说真话规则**（Vickrey's truth serum）。

我们在第 9 章中曾提到过获取信息的成本问题。在拍卖中也不例外。买方在使用维克里方案的拍卖中会泄露其真实的估价，这是因为如此行为会让买方获得利润，而将减少卖方的利润，如同在第一高价拍卖中递减投标价格一样。对卖方而言，这两种方式的优劣根据在哪种方式下利润会递减较多而定。我们将稍后于 16.6 节讨论这一问题，现在，我们先解释维克里方案起作用的原理。

假设你是一个古陶器的收藏者，正在某处的拍卖会中投标一套 19 世纪的茶具，拍卖是以密封式第二高价的方式进行。由于你对古陶器很了解并且你的收藏中恰好缺这么一套茶具，你对该茶具的估价为 3 000 美元，但是你并不知道其他投标者的估价。如果他们不了解古陶器，他们就可能不明确其巨大价值；如果他们对它有情感上的迷恋，他们的估价就会比你估计的价值高得多。

拍卖的规则允许你对此套茶具以任何价格投标，我们将你的投标价记为 b 美元，并考虑其所有可能的值。由于没有小额投标值的限制，我们无法针对此博弈列出有限的支付矩阵，但是我们能够从逻辑上推导出最佳投标价。

显然，你投标成功与否有赖于其他投标者的出价，这是因为你必须考虑你是否会中标。虽然投标结果要视所有其他对手而定，但是只有其中最高出价者才能影响你的结果。我们将这一最高投标价记为 r 美元，并忽略其他低于 r 美元的投标价。

你的最佳投标价 b 为多少呢？你的真实估价是 3 000 美元，我们考察一下比 3 000 美元高或低的投标价，看它们是否会比 3 000 美元的投标价给你带来更好的结果。

我们从 $b>3\,000$ 开始。有三种情况需要考虑：（1）若你对手的投标价低于

3 000 美元（$r<3\ 000$），则你就能以价格 r 美元获得该套茶具，你的利润就是你的投标价格与你的真实估价之差，即（$3\ 000-r$）美元。这同样也是当你的投标价为 3 000 美元时所获得的结果。(2) 若你对手的投标价界于你的实际投标价与你的真实估价之间（$3\ 000<r<b$），你如果仍想赢得茶具的话，将不得不出高于你对它的估价的投标价。此时，你最好出价 3 000 美元；虽然你无法获得该套茶具，但是你也不会损失（$r-3\ 000$）美元。(3) 若你对手的投标价高于你的投标价（$b<r$），你还是无法得到茶具；然而，在此情况下，即使你以真实估价来投标，也是无法获胜的。因此，归纳上述三种情况，以真实估价投标永远不会是差的方式，有时候还会比出价过高来得好。

将投标价递减至 $b<3\ 000$ 时可能的结果又会如何呢？同样地，这里也有三种情况：(1) 若你对手的投标价比你的低（$r<b$），那么你就是出价最高的投标者，则你会以 r 美元获得茶具。你亦可以投标 3 000 美元而得到同样的结果。(2) 若你对手的投标价界于 3 000 美元与你的实际投标价之间（$b<r<3\ 000$），则你的对手会获得茶具。在这种情况下，若你投标 3 000 美元，则你会获得茶具，支付 r 美元，仍然可以获得（$3\ 000-r$）美元的利润。(3) 若你对手的投标价高于 3 000 美元（$r>3\ 000$），那么你就无法获得茶具。在这种情况下，你即使以 3 000 美元来投标，还是无法获得投标物，所以你就算投标 3 000 美元也无妨。因此，以真实估价来投标总不会不是好办法，有时候比起投标过低来得好多了。

如果说诚实地投标永远不会是太差的方法，有时候还会比以低于或高于真实估价的价格投标来得好，那么你最好选择诚实地投标。也就是说，无论你的对手如何出价，诚实总会是上策。换言之，无论是在连续的投标中，还是在离散的投标中，你的占优策略都是以真实估价来投标。

根据维克里说真话规则，在第二高价密封投标的拍卖中，以诚实投标作为占优策略，还有许多其他的应用。例如，若团体中的每一位成员都被问到愿意对有益于整个团体的公共计划资助多少钱时，每个人都会尽量少报自己愿意贡献的数目——让自己成为占其他人贡献便宜的“搭便车者”。我们已经在第 12 章的集体行动博弈中讨论过这样的实例了。维克里方案的变形在这些博弈中也可以引出真实想法。

16.5 全部投标者都要支付的拍卖

在 16.1 节中，我们已经讨论过大部分标准的拍卖类型，但是还没有讨论过可能发生的比较有创意的情况。这里我们要讨论的是这样一种一般价值第一高价密封投标拍卖，其中，所有的投标者，无论是输还是赢，都要支付其投标价给拍卖者。这种连输家也要支付投标价的拍卖似乎不太常见，但是实际上，许多竞赛都会导致这类结果。在政治角逐中，所有候选人都花费了大量金钱和时间，付出了巨大的努力，来筹募经费和展开竞选运动。落选的候选人并不会得到退款。同样地，数百名体育选手花费四年的时间来准备奥林匹克运动会。但仅有一位能够

得到金牌、名声与大家的认可，还有其他两位能够得到银牌与铜牌；所有其余选手的努力都白费了。一旦你开始沿着这条思路思考，你就会了解到何为全部投标者都要支付的拍卖；实际生活中，全部“投标者”都要支付的情形比只有胜者支付的标准的正式“拍卖”多得多。

在全部投标者都要支付的拍卖（all-pay auction）中，你应该如何投标呢？（也就是说，对于时间、努力与金钱的花费，你应采取什么策略？）一旦你决定参加这类拍卖，除非你获胜，不然你的投标就是白费，所以你要有强烈的动机，非常积极地投标。实验中，所有投标的总和通常都大大超过拍卖品的价值，故拍卖者能够获得极大的利润。[7]在每个人都很积极投标的情况下，这还不算是均衡的结果；比较明智的做法似乎就是置身于这种毁灭性竞争之外。但是如果每个人都如此做的话，那么就会有一位投标者，几乎不需要竞争就能够得到拍卖品；这种没有投标的情况也不能算是均衡结果。上述分析让我们知道均衡存在于混合策略之中。

考虑一场有 n 个投标者的拍卖。为使标记简单，我们使用测度单位来衡量，设一般价值之拍卖品的价格为 1。投标价若高于 1 一定有损失，所以我们将投标价限制在 0 至 1 之间。让投标价成为连续的变量 x，x 可以是（0，1）范围内的任何（实）数。由于均衡存在于混合策略中，因此每个人的投标价 x 将会是连续的随机变量。只有当其他投标者的出价都比你的低时，你才会赢得拍卖品。我们将你的均衡混合策略以 $P(x)$ 表示，代表你投标价低于 x 的概率。例如，$P(1/2)=0.25$，表示你的均衡策略是，在 1/4 的时间里以低于 1/2 的价格进行投标（并在 3/4 的时间里以高于 1/2 的价格进行投标）。[8]

照例，我们可通过无差异条件找出混合策略均衡。当其他投标者都选择均衡的混合策略时，任一特定的 x 值对于每位投标者来说必须是无差异的。假设你是 n 位投标者其中之一，投标价为 x。当其他（$n-1$）位投标者的投标价都低于 x 时，你就赢了。其他任何一人的投标价低于 x 的概率为 $P(x)$，任何两个投标价都低于 x 的概率为 $P(x)P(x)$，或 $[P(x)]^2$，其他所有（$n-1$）人的投标价都低于 x 的概率为 $P(x)P(x)\cdots P(x)$（相乘 $n-1$ 次），或为 $[P(x)]^{n-1}$。因此，你赢得拍卖品获得收益 1 的概率为 $[P(x)]^{n-1}$。请注意，无论什么情况下，你都要支付 x。因此，你对投标价 x 的期望支付净值为 $[P(x)]^{n-1}-x$。但是如果你以 0 投标，你就一定会得到 0 支付。由于任何 x 的选择必须是无差异的，这也包括选择 0，因此定义均衡的条件是 $[P(x)]^{n-1}-x=0$。在混合策略均衡中，对于所有的 x 而言，此条件必须成立。所以，混合策略均衡的投标价为

$$P(x)=x^{1/(n-1)}$$

下面举两个实例说明此式代表的意义。首先，请试想 $n=2$ 的情况。对于所有的 x 而言，$P(x)=x$。因此，投标价在已知的 x_1 和 x_2 之间概率为 $P(x_2)-P(x_1)=x_2-x_1$。由于投标价位于此范围内的概率就等于该范围的长度，因此每个投标价都是一样地可能。也就是说，你的混合策略均衡的投标价应该是在 0 至 1 之间随机均匀分布。

其次，我们令 $n=3$，则 $P(x)=\sqrt{x}$。在 $x=1/4$ 的情况下，$P(x)=1/2$，所以

投标价为 1/4 或更低的概率为 1/2。投标价不再服从位于 0 至 1 间的均匀随机分布，其更可能分布于此范围内数值较小的一端。

随着 n 的取值的增大，这种趋势会愈加明显。例如，若 $n=10$，则 $P(x)=x^{1/9}$，且当 $x=1/2^9=1/512=0.001\ 95$ 时，$P(x)=1/2$。在此情况下，你的投标价可能要低于 0.001 95，界于 0.001 95 至 0 之间。因此你的投标价会很接近于 0。

n 越大，你的平均投标价应该会越小。事实上，准确的数学计算显示，若每个人都依此策略进行投标，每位投标者的平均或期望价会刚好为 $1/n$。[9] 当 n 位投标者进行投标时，平均每人的投标价为 $1/n$，所以总投标价为 1，而拍卖者所得的期望利润为零。这表明使用均衡策略可避免过度投标。

当投标者众多时，你的投标价应该更接近于 0 的观点，在直觉上是很合理的；使用均衡投标策略可避免过度投标的发现，使得理论分析家对上述观点更有信心。不幸的是，处于全部投标者都要支付的拍卖之中的人们，并不知道或者忘记了此理论，从而导致过度投标。

16.6 如何在拍卖中进行销售

投标者并不是拍卖中唯一需要考虑最佳策略的参与人。拍卖是一种连续性博弈，第一步要设定规则，第二步才开始进行投标。卖方选择特定的投标方式之后，拍卖结构就随之确定了。

拍卖中，若你想要卖出你的艺术收藏品或房屋，你要决定将使用的最佳拍卖方式。显然，为确保你能从交易中获得最大利润，在下决定前你必须先考虑到各类型拍卖将产生的结果。很多卖方所担心的是，投标者会以低于卖方给出的价格获得拍卖品。针对这一点，大部分卖方都坚持设定拍卖品的**保留价格**（reserve price）。若投标者的出价低于保留价格，卖方则有权从交易中撤回拍卖品。

此外，卖方如何决定拍卖方式以获得最大的净利润？一种可能方式是使用维克里的建议方案，即第二高价密封投标拍卖。根据他的结论，这类拍卖会引出潜在买方的真实投标价。从卖方的角度出发，此作用会让交易变得对其有利吗？

就某种意义来说，在第二高价的拍卖中，卖方会给予投标者适当的利润空间以打消投标者递减投标价以获得高额利润的企图。然而，此举动会减少卖方的收益，类似于在第一高价密封投标拍卖中递减投标价所造成的结果。哪种类型的拍卖最终会对卖方有利，其实是根据投标者对拍卖品风险的态度与对拍卖品价值的信念而定的。

16.6.1 风险中性的投标者与独立估计

当投标者为风险中性（不规避风险）者，且当投标者间各自独立估计拍卖品

的价值时，投标者对风险的信念与态度的结构的复杂性是最小的。第 5 章的附录曾提到，风险中性的个人只关心结果的期望货币价值，而不管这些结果的不确定性程度。估计的独立性是指，一位投标者在估计拍卖品的价值时，不会受到其他投标者估计的影响；该投标者会独立地决定究竟此拍卖品对他的价值如何。在这种情形中，不存在赢家的诅咒问题。倘若维持这些对投标者的假定，则卖方通过四种主要的拍卖类型的任一类型均可期望获得同样的平均收益（在多次尝试后）：英式拍卖、荷兰式拍卖、第一高价密封投标拍卖与第二高价密封投标拍卖。

收益等价（revenue equivalence）并不表示所有拍卖中，每一件拍卖品都能产生同样的收益，而是说，经过多场拍卖后，平均而言会产生相同的售价。在第二高价密封投标拍卖与英式拍卖中，我们能够很容易地见到这样的收益等价。在第二高价密封投标拍卖中，每位投标者的占优策略就是以其真实的估价投标。出价最高的投标者以次高价获得拍卖品，而卖方所得到的价格则等于出价次高者的真实估价。同样地，在英式拍卖中，当价格超过投标者的估价时，他就会退出拍卖，而拍卖一直进行到只剩下最后两位投标者（真实估价最高和次高者）。当价格增加到次高估价时，次高估价的投标者就会退出，此时，最高估价的投标者只需多出 1 分钱就能得到拍卖品。再次强调，卖方所得到的价格就是出价次高者的真实估价。

运用高等数学技巧可以证明收益等价的思想能扩展到荷兰式拍卖与第一高价密封投标拍卖中。所有四种类型的拍卖中，由于投标者为风险中性者，最高估价的投标者会赢得拍卖，而所支付的平均价格等于第二高的真实估价。若卖方想要重复使用某一特定类型的拍卖，那他就不需要过度考虑欲选择的拍卖结构，因为这四种类型的拍卖都会产生相同的期望价格。

在实际拍卖中，实验性证据已经验证了收益等价定理的有效性。实验结果显示，对相同的拍卖品与投标者而言，平均来看，荷兰式拍卖价比第一高价密封投标拍卖价更低，这可能是由于荷兰式拍卖存在某种与不确定性因素相关的正效用。这些实验也表明，过高出价（出价高于你所知道的价值）出现在第二高价密封投标拍卖中，但是并不出现在英式拍卖中。这一行为说明，当投标者必须明确某种价格时，他们会叫价更高，就如同在密封式拍卖中那样，在这类拍卖中投标者似乎更看重投标价与最终获取拍卖品的概率之间的关系。而基于互联网拍卖的实际证据却发现了截然相反的结果，荷兰式拍卖平均收益比第一高价密封投标拍卖高出 30%。荷兰式拍卖中额外的投标者利益或在长达 5 天的拍卖中的不耐烦可以解释这一反常现象。基于互联网拍卖的实际证据还发现其他两种类型的拍卖也存在收益等价。

16.6.2 风险规避的投标者

我们继续假设投标价与信念无关，而拍卖结果却受投标者对风险的态度影响。特别是假设在投标者为风险规避者的情况下，他们更关心出价过低导致的损失——失去拍卖品，而较少关心与投标相关的成本或他们的真实估价。因此，

一般而言，如果可能，风险规避的投标者尽量想在没有过度投标的情况下赢得拍卖品。

在第一高价与第二高价（密封投标）拍卖中，各自的投标结构会产生什么影响？请将第一高价的拍卖等同于荷兰式拍卖。这里，风险规避的特性使得投标者宁可早一点投标而不愿意晚一点投标。当投标价落在略高于投标者的估价处时，等待投标变得越来越有风险。对于风险规避的投标者，我们预期他们会很快地投标，而不会因为想要多赚一点利润而稍加等待。将此道理应用于第一高价密封投标拍卖中，我们预期他们投标价递减的幅度会小于他们不是风险规避的投标者时其投标价递减的幅度——太大幅度地递减价格会增加失去拍卖品的风险，这正是风险规避的投标者想要避免的。

将此结果与第二高价拍卖（获胜的投标者支付次高的投标价）的结果相比较。投标者在第二高价拍卖中以其真实估价进行投标，而支付较少的价格。若他们只在第一高价拍卖中稍微降低其投标价，则投标价会趋近于投标者的真实估价，而投标者将支付其投标价。虽然投标价将会下降一点，但是在第一高价拍卖中最后支付的价格可能会高于在第二高价拍卖中支付的价格。因此，当投标者是风险规避者时，卖方最好选择第一高价拍卖，而不是第二高价拍卖。

在风险规避与密封投标的情况下，卖方最好使用第一高价拍卖。若拍卖形式为公开喊价（英式），那么投标者对风险的态度与结果无关。因此，在这种情况下，对于卖方而言，风险规避并不会改变结果。

16.6.3 相关的估计

假设在决定对拍卖品的真实估价时，投标者会受到其他投标者的估价的影响（或是受到估价信念的影响）。这种情况与一般价值的拍卖是相关的，例如 16.2 节中的油气田勘探实例。假设你的专家尚未给出关于油气田未来利润的明确的测算结果，你正对其潜在利润感到悲观，于是你得出了能够反映你悲观感觉的估计值 V。

在这种情况下，你可能会猜想，你的对手或许也从他们的专家那儿得到了负面的报告。当投标者相信他们的估价都相似时，比如都同样相对地高或低，我们称这些对价值的信念或估计为正相关。请回忆我们在第 9 章附录中介绍的相关性概念，那里我们讨论的是相关风险与保险。在这里，你的对手的估价也不高的可能性，更扩大了悲观对你做评价时的影响。若你参与的是第一高价密封投标拍卖，此时相比于投标者的估价相互独立时，你降低投标价的幅度会更大。当然，若投标者对油气田的利润前景持乐观态度，对其的估价一般也较高，此时与投标者的估价相互独立时相比，投标者估价间的相关性会使得投标价下降的幅度较小。

然而，在第一高价密封投标拍卖中，随着低价（或悲观的）投标而来的投标价递减，对于卖方而言是个警醒。由于投标者信念的正相关，卖方可能想避免使用第一高价密封投标的拍卖方式，而想使用维克里所推崇的第二高价密封投标的

拍卖方式。我们刚刚才谈到此类型的拍卖能够鼓励诚实投标，当投标者对拍卖品价值的估计存在相关性时，卖方通过避免采取能导致投标者降低投标价的拍卖方式可获得更大的利益。

英式拍卖的最后结果与第二高价密封投标拍卖的结果一样，荷兰式拍卖的结果与第一高价密封投标拍卖的结果相同。因此，当面对一群对投标品价值的估计存在相关性的投标者时，卖方应该偏好英式的公开喊价拍卖，而不是荷兰式的公开喊价拍卖。假设你在公开喊价的英式拍卖中竞标一块油田，目前投标价非常逼近你的估计值，而你的对手仍然狂热地投标着，你可以推测他们对油田价值的估计值至少和你的一样高——可能比你的还高很多。观察对手投标行为而得到的信息可能会让你相信，你的估计价格太低。因此，你可能在投标过程中提高你对油田的估计价格。而你持续投标的行为可能会成为其他投标者继续投标的动力，这一过程还会再持续一阵。如果这样的话，卖方会得到利润。更普遍的是，当投标者间的估计是相关的时，卖方能够预测，英式拍卖中的销售价格将高于第一高价密封投标中的销售价格。然而，对于投标者而言，公开投标的作用则是传递额外的信息及降低赢家的诅咒。

以上在假设有众多投标者参与的情况下对相关估计进行了讨论。然而，若只有两个人参与拍卖，则公开喊价的英式拍卖有利于卖方，因为投标者都会特别狂热地竞标该标的物。他们会尽可能地持续相互竞标，一直在较低估价的基础上提高价格，所以双方的价格比起价高很多。然而，当其中一位投标者的估价很低时，同样的拍卖方式就可能对卖方很不利，因为另一位投标者的投标价格很可能大大高于前者的投标价格。在这种情况下，我们称此两位投标者的估价负相关。当买方人数很少且估价差异可能很大时，我们建议卖方选择荷兰式拍卖或第一高价密封投标拍卖的方式。这两种方式都可以降低高估价的投标者以低于其真实价格赢得拍卖品的可能性，也就是说，这两种方式将利润从买方转移给卖方。

16.7 其他需要思考的问题

16.7.1 多个投标品

当考虑拍卖多件物品时，例如银行拍卖未如期偿还贷款的车辆，或房地产销售人员拍卖房屋，你可能想象拍卖者将拍卖品一件件拿到拍卖台上，然后分别卖给出价高的投标者。当每位投标者对不同拍卖品的价值有各自独立的估计时，这一过程是正确的。然而，对投标者的估价进行建模时，假设投标者对拍卖品价值的估计是独立的，并非总是合适的。若投标者对整套拍卖品的估价，高于其对所有单件拍卖品的估价之和（整套包括许多单件），则逐件拍卖或一次性拍卖将会有不同的投标策略，也会有不同的结果。

试想，一家名为瑞德的房地产开发商想购买一大片土地，建造一处针对专业人士的住宅社区。C与M两个小镇拍卖的土地都符合其要求。两块土地都是4英亩的正方形土地。C镇镇长指示拍卖者，一次卖出一小块面积为1/4英亩的土地，由边缘向中间将地卖出，先卖出角落的土地，然后再依次卖出北边、南边、东边与西边的土地。同时，M镇镇长指示拍卖者一次性拍卖一整块面积为4英亩的土地。但是，如果投标者没有超过保留价格，那就将土地分成两块较小的面积为2英亩的土地分别拍卖；若还是没有售出，再将土地分成更小的四块面积为1英亩的土地分别拍卖。

根据大规模的市场分析，对于瑞德而言，C镇和M镇上的两大块土地的价值是相同的。然而，它必须得到任何一镇里4英亩的土地才会有足够的空间建造新住宅区。两个镇里的拍卖都在同一天举行。它应该参加哪一场呢？

我们清楚到看到，在M镇会比在C镇更有机会以合理价格——低于或等于它真实估价的价格——购得4英亩的土地。在M镇的拍卖中，它只要观察拍卖进行的过程，当第二高价仍低于他的真实估价时，就以最高价格投标。而在C镇的拍卖中，它必须赢得标售16块土地的每一场拍卖才能获得4英亩的土地。在这种情况下，它必须估计到，拍卖过程中，当小块土地数目不断减少时，对C镇土地有兴趣者就都会变得更急切想要得到该镇的土地——甚至会使用暴力。当瑞德在前几场小块土地的拍卖中激烈地出价时，它同时也要保守计算以确定，在拍卖结束前，其所有的投标价格之和将不会超过它对所有小块土地的真实估价的总和。在这种拍卖中设计投标策略是很不容易的，而且在有利润的状况下获得所有土地的可能性不大，因此瑞德偏好参加M镇的拍卖。

请注意，站在卖方的角度，在有相当数量的投标者对小块土地有兴趣的情况下，C镇的拍卖方式会比M镇的拍卖方式获得更多利润。但是，若只有像瑞德一样的开发商参与土地的拍卖，那么这些开发商甚至会犹豫是否要参加C镇的拍卖，因为它们担心会因在任何仅仅一场小块土地的拍卖上的失误而遭受重大损失。在这种情况下，M镇的拍卖就比较有利于卖方。

C镇可以通过修改拍卖规则以消除房屋开发商的担忧。特别是它不必逐块拍卖每块土地，而是采用某种可让所有土地同时被获得的简单的拍卖方式。下面的拍卖方式是可行的，它可使各个投标者能够确定其所需要的土地数目以及愿意对每块土地所支付的价格。出价总额最高的投标者将赢得所希望数目的土地，而这种出价总额等于所需要的土地数目与每块土地单价的乘积。如果在中标者获得其所需要的土地后还有土地剩余，那么，其余土地可以按类似方式出售，直到所有土地被售完。这种机制会给对大片土地感兴趣的投标者中标机会，使其为了数块土地而相互潜在地进行出价竞争。因此，C镇可能发现这种拍卖最终更有利可图。

16.7.2 系统缺陷

前面我们讨论了在已知投标者对风险的态度与他们的估计是否相关的条件

下，哪种拍卖结构最有利于卖方。然而，投标者永远都想利用投标策略击败卖方。一项设计完善且能够获利的拍卖计划，总是会被一位——更多的是一群——同等聪明的投标者所击败。

若一场拍卖只有很少的投标者参加，他们就能够通过彼此勾结击败卖方，就算是卖方采用维克里的第二高价密封投标方式仍会如此。互相勾结的投标者会提出一个高价（第一高价）与一个很低的第二高价，然后他们就能够以很低的次高价获得拍卖品。此结果成立的条件是，必须没有其他投标者以中间价投标，或是此勾结团体必须要有能力防止其他人以中间价格投标。虽然在本例中，保留价格只抵消了卖方的一部分损失，然而勾结的可能性凸显了卖方设置保留价格的必要性。

第一高价密封投标拍卖较能防止投标者互相勾结，理由如下。当投标者都不想诚实投标时，潜在的勾结团体会陷入多人的囚徒困境博弈，在此博弈中，每个投标者都有欺骗的动机，在这种动机的驱使下，他们会以高价投标赢得拍卖品而违反协定不与勾结团体的其他成员分享利润。因此，在这种拍卖中，很难维持投标者间的勾结，因为欺骗行为很容易发生（也就是说，所提出的投标价不同于勾结团体所协议的价格），但是被欺骗者却很难察觉。因此，密封投标拍卖的特性就是，一直要到拍卖结果公布后，才能察觉到欺骗行为，也才能处罚欺骗者，但是已经为时太晚。然而，若某团体连续参与数场类似的拍卖，他们将有更多机会进行勾结，从而在重复的博弈中达成均衡。

可以根据投标者个人或团体在特定拍卖中的需要来设计投标方案。一个非常聪明的例子是美国 FCC 拍卖美国电磁波频段的实例，特别是个人移动电话服务频段的拍卖。在观察到前几场拍卖中价格不断飙升后，投标者明显地渴望压低中标价。其中三家公司（后来受到联邦司法部起诉）使用 FCC 代码或特定地区的电视区号作为其投标价的最后三位，通过这种方式传递其真实的意图，以获得这些特定地区的执照。FCC 已宣称这种做法显著地降低了这些地区的执照的最后价格。此外，在此之前的几场宽带拍卖中，显然它们都使用了其他传递信号的装置。有些厂商明白地表示它们想获得特定的执照，有些厂商则运用各种策略性投标技巧，向大家表示它们在特定执照上的利益，或劝阻其他对该种执照也有兴趣的厂商的投标行动。例如，在一场宽带拍卖中，GTE 和其他厂商都明显地使用了代码投标技巧——其投标价以可以通过按键式电话的数字键显示其名字的数字结尾。

这里我们简要提出一点，即欺骗行为并不仅仅限于拍卖中的投标者，卖方也可能使用狡诈的做法去抬高他们想拍卖的物品的最后价格。比如，当卖方能够在自己的拍卖中虚假竞标（做手脚）时，“托”的欺骗（shilling）就会发生。只有在英式拍卖中，骗子才有可能雇用一个为卖方工作的代理人，他假装是一个平常的拍卖投标人。在互联网拍卖站上，欺骗实际上很容易进行，因为卖方可以在自己的拍卖中再登记一个人，并让此人参与投标。所有网上拍卖都有规则与监督机制以阻止这种行为。如果第二高价密封投标拍卖的卖者（不公开地）提高第二高价水平，他们也能获利。

16.7.3 信息泄露

最后，我们来考虑这样一种可能性——卖方拥有关于拍卖品的某些私人信息，这些信息能够影响投标者对拍卖品的估价。当买方很重视拍卖品的品质或耐用性时（例如汽车、房屋或电器），就会产生这种情况。卖方过去关于此类物品的经验，是获胜的投标方的未来利润的一个很好的预测指标。

在这种情况下，卖方必须仔细考虑任何隐藏信息的企图。当投标者知道卖方拥有某些私人信息时，如果拍卖者未能披露这些信息，投标者就可能会将之视为表明这些私人信息是不利的的一个信号。即使卖方的信息是不利的，他最好是披露这些信息，因为投标者对这些信息的评价可能比真实信息更糟。因此，诚实总是上策。

诚实有利于卖方，还有另外一个原因。当卖方拥有一般价值物品的私人信息时，他应该披露信息，以提升投标者对物品的估价。投标者对估价的正确性愈有信心，则他们的投标价愈有可能高于其估价。因此在一般价值的拍卖中，披露卖方的私人信息，不仅有助于卖方减小投标者的价格递减幅度，而且也有助于买方降低赢家的诅咒。

16.8 互联网上的拍卖

拍卖已经成为世界上电子商务的重要内容。每个月都有 4 000 万人参与网上拍卖，每年交易额达到 600 亿美元。拍卖站点本身可以通过列表收费、卖方代理费和广告收益赚钱，每年可达 10 亿到 20 亿美元。随着网上拍卖交易量与收益逐年增加，电子商务作为一个特殊领域激起了许多工业分析家、消费者和拍卖理论家浓厚的兴趣。

网上拍卖站点存在不足十年。在 Onsale. com 1995 年 5 月出现后不久[10]，1995 年 9 月 eBay 站点开始运行。现在已有大量拍卖站点——100～150 个不同的有效站点，同时由于新站点不断出现，现存有些站点进行合并以及无利润的小站点关闭，因此站点的准确数目在不断变化。这些站点，无论大小，都以很多方式出售大量物品。

在一些较大的被较多地使用的网站上，如 eBay 和 uBid，大多数物品都被分类归入“收藏品”，它们可以是从一件古董明信片到一个塑料碗的任何东西。但是也有一些特殊的拍卖网站处理邮票、酒和烟、警察从搜捕中取得的财产、医疗设备、较大的建筑设备。无论什么类型的站点，其拍卖的大部分物品都被认为是用过的，这样，消费者只需用手指点击，就有机会进入世界上最大的旧货销售地。

关于出售物品的信息与文献中的一个假设是吻合的。这一假设认为对于销

售数量有限的物品而言，网上拍卖是最有用的，因为对这些物品的需求是未知的，且卖方也不容易决定一个适当的价格。拍卖过程能够有效地找出这些物品的"市场价格"。这些物品的出售方在网上可以有最好的获利机会，那里有大量旁观者能够提供以往未知的需求参数。消费者也能够获得渴望获得但不了解的物品，同时还可能获得利润。经济学家们称此为匹配过程，在这个过程中，每个行为人都从交易中取得正利润，从而交易是有效率的。（有效机制会在第 18 章得到深入讨论。）

除了出售许多不同类型的商品外，网上拍卖还制定了多种拍卖规则。很多网站实际上提供了若干种拍卖类型，并当一个卖方打算出售一件物品时，允许他选择自己的拍卖规则。最普遍使用的规则是英式拍卖和第二高价密封投标拍卖。绝大多数拍卖站点都提供一种或两种拍卖方式。

提供真正英式拍卖的若干网站获悉高价后，会将其公布，并在拍卖结束时，获胜者支付其叫价。此外，还有很多网站使用英式拍卖形式但是允许进行所谓的**代理投标**（proxy-bidding）。代理投标实质上使得英式拍卖变成了第二高价密封投标拍卖。

当进行代理投标时，投标者申报他愿意支付某件物品的最高价（**保留价格**）。买者不需出示最高价格，而是在最接近的高价上，拍卖网站显示一个单位叫价增量，然后代理拍卖系统为买者出价，每次高于其他投标者一个单位叫价增量，直到达到买方申报的最高价格为止。这个系统允许拍卖获胜者支付比第二高价多一个单位增量的叫价，而非支付他自己的叫价。大多数大拍卖网站包括 eBay、uBid、Amazon 和 Yahoo，都提供代理投标。此外，一些网站还提供明确的第二高价拍卖。在此拍卖形式下，实际的最高价是要公布的，但是获胜方按照第二高价支付价格。

荷兰式拍卖网站很少。有文献报道，一些采用降价拍卖的网站已经不复存在；只有一些零售站点还在提供相当于荷兰式拍卖的拍卖，如 Land's End 每周末在自己网站的某个特别的地方公布一些过多的存货——这些物品的价格会在下周降低三次——并在周末撤下未被购买的物品。CNET Shopper 会公布电子产品的最近降价情况，但不保证其他产品在某个特定时期内会降价。人们猜测荷兰式拍卖的缺陷在于具有某种即时的特征，从而使得这些拍卖很难在网上展开。

然而，仍有一些站点提供荷兰式拍卖，这些拍卖和与其配套的拍卖类型**扬基拍卖**（Yankee auctions）实际上就是在单场拍卖中提供多个（相同）单位拍卖品，类似于 16.7 节中的 C 镇的土地拍卖。这种拍卖提供给投标者对一个或多个单位拍卖品出价的选择。术语"扬基拍卖"其实就是我们所叙述的城镇土地拍卖这种拍卖。投标总价（含价格与数量）最高的投标者赢得物品而且投标者按各个单位拍卖品的投标价进行支付。"荷兰式拍卖"这一名称被保留下来，用以指称这样一种拍卖，其中投标价按总额排序，拍卖结束时，所有中标者按最低的中标价进行支付。

网上拍卖与现实拍卖相比较，规则与结果十分相似，前两节所论述的策略性问题在网上拍卖中也存在。网上拍卖对于买方而言是好的，因为他们容易"参与"，而且为了找到自己感兴趣的商品，只需要简单地使用搜索引擎；同样地，

卖方也非常容易找到大量的买者，且能够较方便地选择自己偏好的拍卖规则。另一方面，网上拍卖也有欠妥的一面，由于在投标前买方不能检查物品，因此买方与卖方必须相信对方会按照既定的承诺进行支付或递送。

然而，现实拍卖与网上拍卖最大不同在于拍卖结束方式，现实中（英式或荷兰式）拍卖在没有其他投标者叫价时停止，网上拍卖则需要制定一个特殊的结束规则。在几乎所有的现实情形下，网上拍卖都在经过一段特定的时间（通常为7天）之后结束。然而，不同站点使用不同结束规则，一些站点如 Amazon 允许在结束时限到达之前10分钟内出现新叫价时延长结束时间；其他站点如 eBay 则不管在时限到达时出现多少叫价，都会结束拍卖。

两种结束时限规则下投标者的不同行为引起了研究者浓厚的兴趣。阿文一罗斯和阿克思一俄肯菲尔收集的证据表明：eBay 的硬性结束时间方式使得投标者可因迟投标获利。这种行为被称为“阻击”，出现在私人价值和公共价值拍卖中。

从策略上看，eBay 网上晚投标者可通过避免同其他投标者——这些投标者不使用代理投标系统，在整个拍卖过程中也不更新自己的投标价——发生投标战而获益。此外，通过保护好他们关于商品的一般价值的私人信息，晚投标者也可以获利。即使在 eBay 网上，经常使用其提供的服务的人也能辨别出哪些投标者知道特别好的拍卖品（尤其是古物）。具有这种信息的投标者可能想对别人隐藏信息，以确保最后能压低销售价。这样，由于购买某些物品如古董时需要买方具有专门的鉴别技巧，因此买方可能偏向 eBay 而不是 Amazon。同样道理，这些物品的卖方可以在 Amazon 得到更好的价格——但仅当有经验的买方愿意在那儿投标时。证据表明，eBay 的重复性投标者趋向于在更接近结束时限时叫价；Amazon 的投标者则明白他们可以更早叫价。这个发现可以证明朝向均衡的叫价行为以相同方式变动，投标者在不同拍卖中的策略与我们关于最大化支付的行为的预期是一致的。

16.9 拍卖理论与实践的其他读物

很多关于拍卖理论的文献应用到相当复杂的数学知识。一些关于拍卖行为与结果的富有见地的内容可以在如下文献中找到：Paul Milgrom 的 *Auctions and Bidding: A Primer*、Orley Ashenfelter 的 *How Auctions Work for Wine and Art* 和 John G. Riley 的 *Expected Revenues from Open and Sealed Bid Auctions*。这些论文都刊登在 *Journal of Economic Perspectives*，vol. 3，no. 3，1989，pp. 3－50。这些论文都是可阅读的，你可以获得有关计算方面的丰富的背景知识。

关于本主题的更多的复杂信息也可以得到。R. Preston McAfee 和 John McMillan 有一篇综述性文章 *Auctions and Bidding*，发表在 *Journal of Economic Literature*，vol. 25，1987，pp. 699－738。最新的文献还有 Paul Klemperer 的 *Auction Theory: A Guide to the Literature*，发表在 *Journal of Economic Surveys*，vol. 13，no. 3，1999，pp. 227－286。这些文献都应用了与拍卖理论相关的

高深的数学知识，同时还给出了综合性的参考文献。

维克里关于第二高价拍卖中真实投标的详细论述见发表在 *Journal of Finance*，vol. 16，no. 1，1961，pp. 8 - 37 的 *Counter speculation，Auctions，and Competitive Sealed Tenders*。这篇文章是最早注意到收益等价存在的文章之一。更多研究拍卖类型收益结果的成果可参见 J. G. Riley and W. F. Samuelson，"Optimal Auctions，" *American Economic Review*，vol. 71，no. 3，1981，pp. 381 - 392。关于维克里第二高价拍卖的可读性强的历史文献介绍可参见 David Lucking-Reiley，"Vickrey Auctions in Practice：From Nineteenth Century Philately to Twenty-First Century E-commerce，" *Journal of Economic Perspectives*，vol. 14，no. 3，2000，pp. 183 - 192。

关于拍卖行为的实验性证据可参见 John H. Kagel，"Auctions：A Survey of Experimental Research，" in John Kagel and Alvin Roth，eds.，*The Handbook of Experimental Economics*，Princeton：Princeton University Press，1995，pp. 501 - 535。关于收益等价的更多研究可参见 David Lucking-Reiley，"Using Field Experiments to Test Equivalence Between Auction Formats：Magic on the Internet，" *American Economic Review*，vol. 89，no. 5，1999，pp. 1063 - 1080。另外一些关于网上拍卖行为的研究可参见 Alvin Roth and Axel Ockenfels，"Last-Minute Bidding and the Rules for Ending Second-Price Auctions：Evidence from eBay and Amazon Auctions on the Internet，" *American Economic Review*，vol. 92，no. 4，2002，pp. 1093 - 1103。

关于美国 FCC 电磁波频段拍卖与拍卖设计的问题，详见 Peter Cramton and HesseSchwartz，"Collusive Bidding：Lessons from the FCC Spectrum Auctions，" *Journal of Regulatory Economics*，vol. 17，2000，pp. 229 - 252；Paul Klemperer，"What Really Matters in Auctions in Auction Design，" *Journal of Economic Perspectives*，vol. 16，no. 1，2002，pp. 169 - 189。关于网上拍卖的文献综述可参见 David Lucking-Reiley，"Auctions on the Internet：What's Being Auctioned，and How?" *Journal of Industrial Economics*，vol. 48，no. 3，2000，pp. 227 - 252。

16.10 总　结

除了标准第一高价公开升价拍卖或英式拍卖外，还有荷兰式降价拍卖，以及第一高价与第二高价密封式拍卖；拍卖品对于每个投标者而言，可能有唯一一个一般价值或诸多不同的私人价值。在一般价值拍卖中，投标者通常仅当他的叫价过高时才能获胜，并陷入赢家的诅咒之中；在私人价值拍卖中，最优投标策略，包括关于何时从自己真实估价出发递减叫价的决策，依赖于采用的拍卖类型。在常见的第一高价拍卖中，存在一种压低叫价的策略动机。

维克里认为卖方可以通过使用第二高价密封式拍卖诱导出投标者的真实信息。一般而言，卖方可以选择拍卖机制以确保自己获得最大利润，这种选择将依

赖于投标者对风险的态度与对拍卖品价值的估计。若投标者是风险中性的，并独立地对拍卖品的价值进行估计，则所有拍卖类型将产生相同的结果。如何拍卖大量物品——单个的或者成套的？是否披露信息？对上述问题做出决定并非易事，卖方必须警惕投标方的勾结与欺骗行为。

互联网是当前物品拍卖出售的最大集散地之一，在线拍卖站点出售许多不同类型的物品并使用许多不同类型的拍卖方式。现实拍卖与网上拍卖的主要不同之处在于某些站点特别是 eBay 设置了硬性结束时间。对这种拍卖结束规则，晚点叫价和阻击行为被证明是理性的反应。

关键术语

所有投标者都要支付的拍卖（all-pay auction）
升价拍卖（ascending auction）
一般价值（common value）
降价拍卖（descending auction）
荷兰式拍卖（dutch auction）
英式拍卖（english auction）
第一高价拍卖（first-price auction）
客观价值（objective value）
公开喊价（open outcry）
私人价值（private value）
代理拍卖（proxy-bidding）
（投标者）保留价格（reservation price）
（卖方）保留价格（reserve price）
收益等价（revenue equivalence）
密封投标（sealed bid）
第二高价拍卖（second-price auction）
价格递减（shading）
欺骗（shilling）
主观价值（subjective value）
维克里说真话规则（Vickrey's truth serum）
赢家的诅咒（winner's curse）
扬基拍卖（yankee auction）

习　题

1. 一名油漆工在替建筑商工作时签有正式合同。一般而言，他对工作成本的估计都是很准确的：有时候会偏高一点，有时会偏低一点，但是平均来说是准确的。在工作淡季时，他会投标以争取其他的工作。他说："这两种工作不同，因为后者的工作成本总是高于我的预期。"假设他的估计技巧并没有因为工作不同而改变，那么如何解释这一差异现象呢？
2. "投标者为风险规避型时，一个在拍卖中出售自己房屋的人，通过采用第一高价密封投标拍卖方式，将获得更高的期望利润。"这一论断是对还是错，请给出你的答案并解释原因。
3. 请试想有一场拍卖要卖出 n 件完全相同的物品，有 $(n+1)$ 位投标者。每件物品的实际价值都是一样的，对于所有投标者而言也都相同；存在误差的情况下，每位投标者对物品的一般价值只有一个独立的估计值。这场拍卖采用密封投标方式。出价最高的 n 位投标者，每人各得到一件拍卖品，并支付其

投标价。在这种情况下，什么因素会影响投标者的投标策略？怎样影响？

4. 假设有三位风险中性的投标者，皆有兴趣购买一顶王妃无檐小帽。投标者（编号从1到3）分别估价12美元、14美元、16美元。她们将以如（a）～（d）所描述的方式参与竞拍。在每种情况下，投标者可以在5美元到25美元任何价格上以1美元的增量进行投标。
   (a) 哪个投标者将在公开喊价英式拍卖中获胜？赢家最后支付的价格为多少？利润为多少？
   (b) 哪个投标者将在第二高价密封投标拍卖中获胜？赢家最后支付的价格为多少？利润为多少？将之与（a）所得到的结果进行比较。在这两种情况下，利润差异的原因是什么？
   (c) 假设采取第一高价密封投标拍卖，所有投标者的出价都低于其真实估价若干正的增量（至少1美元）。这种拍卖可能的结果是什么？将之与（a）和（b）所得的结果进行比较。卖方有任何明确的理由选择这些拍卖机制中的一个吗？
   (d) 在（c）中，风险规避型的投标者将降低其中投标价的递减幅度。假设他们一点都不会递减，（c）中获胜的价格（和投标者的利润）是多少？卖方会关心采用哪种类型的拍卖吗？为什么？

5. 赢家的诅咒可以使用不同于正文中的方法来解释："只有当你中标时，你的投标才有作用；也只有当你的估计值高于其他所有投标者时，你才会赢。因此，你应该注意这一事实。也就是说，你应该把所有其他投标者的估计值都当做比你的估计值低，并使用这一'信息'来修正你本身的估计值。"请将此思想运用到一个非常不同的情况。

   有一个由12人组成的陪审团，他们听到和看到相同的证据，但是他们都根据各自的想法与经验解读这些证据，从而估计被告有罪或无罪。每位陪审员都会被要求投下"有罪"或"无罪"的一票。若陪审团投下12票"有罪"，被告就会被判有罪；若有任何一人或一人以上投票"无罪"，则被告就会被释放。每位陪审员的目标是根据证据做出最准确的裁决。每位陪审员都会采取策略性投票，并应用所有我们所学过的信息推断的方法，企图极大化其目标。请问他们投票的均衡结果会达到最佳状况吗？

6. 你现在在二手车市场中，看到了喜欢车型的广告。车主并没有设定价格，但欢迎有兴趣的买家出价。你在购买前的市场调查只让你对其价格有了一个粗略的概念：你觉得价格可能界于1 000至5 000美元之间（所以你计算出的平均价格为3 000美元）。目前的车主知道车子的确切价值，若你的出价超过此值，他就接受。当他接受你的出价而你获得车子后，你就会知道车子的真实价值。由于你了解一些修车技术，所以知道一旦获得这辆车子，不论车子当时价值为多少，你会对车子进行修护，从而使其价值提升1/3。
   (a) 若你出价3 000美元，你的估计期望利润为多少？应该以此价格投标吗？
   (b) 在没有损失情况下，你的最高投标价为多少？

7. ［选做］用数学分析投标策略与均衡是很困难的，但是若你已了解一些微积分的基础知识，我们将在此为你提供一个简单例子。

请试想有一场第一高价密封投标拍卖，共有两位投标者参加。每人都知道他自己的估价，但不知道对方的。两者的估价都有可能在 0 至 1 间的任何位置上，也就是说，每位投标者的估价分布于 [0，1]，每个人的策略能被预先表示成完整的行动方案："若我的估价为 v，我将会出 b。"换言之，此策略将投标表达成估价的函数，$b=B(v)$。我们可在此策略中寻找到均衡。参与人的策略是完全相同的，而我们要在他们相同的策略（B 函数）中，找到对称的均衡。当然，我们必须利用均衡条件来求解 B 函数。

假设另一位参与人正在使用 B 函数。你的估价为 v，而你正在考虑你的最佳行动。你可以将你的估价视为 x，而投标价为 $B(x)$。对手的估价低于 x 时，你将以此投标价获胜。由于所有在 [0，1] 中的估价可能性都相同，对方估价比 x 小的几率是 x。如果你赢了，你的剩余就是你的真实估价减去你的投标价，也就是 $v-B(x)$。因此，如果你的价值为 x，你由投标策略得到的期望支付就是 $x\,[v-B(x)]$。

(a) 请写下能最大化你期望支付的关于 x 的一阶条件。

(b) 在均衡时，你要遵守 B 函数的策略，也就是，你将选择等于你的真实估价 v 的 x。请使用该信息重新写出 (a) 中 x 的一阶条件。

(c) 将 $B(x)=v/2$ 替换至 (b) 中所得的方程中，并观察其状况。如果 (b) 中的方程式有唯一解，它也就是我们在寻找的均衡策略，请直觉地解释其意义。

8. [选做] 请将习题 7 中两位投标者改为 n 位并重新做一次；请证明 $B(v)=v(n-1)/n$ 能提供所有投标者均衡策略。请直觉地解释为什么当 n 增加时，对于任何 v，$B(v)$ 也会跟着增加？

【注释】

[1] 关于拍卖理论与实践的其他文献可参看本章最后一节。

[2] 例如，美国是把近海石油勘探权以拍卖方式租给公司，包括墨西哥湾和阿拉斯加海岸等地的石油勘探权。宾夕法尼亚州则一直考虑着对其近 50 万英亩森林陆地天然气勘探权进行拍卖。

[3] 这 10 个估计值平均而言会分布在 9 亿至 11 亿美元之间（10 亿的左边或右边）。最高或最低的估计值，平均而言，都会落在分布区域的两个极端值处。

[4] 参见 Steven Landsburg, *The Armchair Economist*, New York: Free Press, 1993, p. 175。

[5] 参见 Richard Thaler, "Anomalies: The Winner's Curse," *Journal of Economic Perspective*, vol. 2, no. 1, 1988, pp. 191 - 201。

[6] 维克里是过去 40 年中在经济学上富有原创性的思考者之一，1996 年他以对拍卖与披露真实估价的过程的研究获得诺贝尔经济学奖。但是不幸的是，他在奖项宣布三天后与世长辞。

[7] 作者之一迪克西特曾在博弈理论课堂中拍卖 10 美元的钞票，并从一个有 20 位学生的小组中获利 60 美元。在普林斯顿大学里有一个传统，在学期末的最后一堂课中学生会持续鼓掌向教授表达敬意。有一次，迪克西特提供 20 美元给鼓掌最久的学生。这是一场公开喊价全部投标者都要支付（以鼓掌形式）的拍卖。当大部分的学生都在 5～20 分

钟后放弃后，仍有三位学生持续鼓掌了 4.5 小时。

[8] $P(x)$ 称为随机变量 x 的累积概率分布函数。对 x 而言，较熟悉的概率密度函数来自对 x 的导数，$P'(x)=p(x)$。$p(x)\mathrm{d}x$ 代表变量位于 x 至 $x+\mathrm{d}x$ 之间的概率。

[9] 每位参与人的期望投标价相当于 x 的期望值，请用概率密度函数 $p(x)$ 来计算。在本例中，$p(x)=P'(x)=(1/n-1)x^{(2-n)/(n-1)}$，而 x 的期望值为 $xp(x)\mathrm{d}x$ 从 0 到 1 的总和 $\int_0^1 xp(x)\mathrm{d}x = 1/n$。

[10] Onsale. com 与 Egghead. com 合并于 1999 年，Amazon 于 2001 年末收购了该合并公司的资产，这三个拍卖站点现在简称“Amazon. com Auctions”。

第 17 章　谈　　判

在日常生活中，处处都涉及谈判。孩子从小就开始和其他小朋友以谈判分享玩具和参与游戏。夫妻之间通过谈判解决房屋归属、养育孩子以及为彼此事业所做的调整等问题。买卖双方谈判价格，劳资双方商议薪金。国家之间谈判自由贸易政策，强权国家之间谈判裁减武器议题。本书两位作者也得谈判——通常是以和平的方式进行——哪些章节要增添内容或省略内容，如何展开本书的论述等主题。在谈判中欲获得好结果，参与人必须设计良好的策略。本章将提出并解释这些思路与策略。

所有谈判都有两个相同点。第一，谈判中双方达成协议所得的总支付应大于双方各自处理事件而得到的支付总和，也即整体必须大于部分的和。若不存在此超额价值（或称为“剩余”），谈判就没有意义。比如两个小孩正在考虑是否一起玩耍，如果一起玩耍无法让他们玩到更多的玩具，或无法玩到对方正在玩的玩具，那么他们最好还是拿着自己的玩具独自玩耍。当然，世界上充满了太多不确定性，可能无法具体化期望利润。然而，假如要从事谈判，就要使参与人至少有“获得”的感觉：当浮士德与撒旦谈判时，浮士德就觉得从撒旦那儿得到的知识与力量，比起他最后必须付出的灵魂更加有价值。

第二个重要的相同点由第一点产生：谈判并非零和博弈。当剩余存在时，谈判就是解决如何分配剩余的问题。每位谈判者都试图多为己方争取一点好处，而给对方少留一点。这看起来或许是零和博弈，但是这存在着一种危险：若无法达成协议，将没有人能够得到剩余。存在对双方都有害的选择，以及双方都试图逃

避这种选择，正是威胁可能存在的原因。这种威胁可能是明确的，也可能是隐含的。不论是明确的还是隐含的，它使谈判成为一种策略性的活动。

在博弈论出现以前，一对一的谈判被视为是困难的，而且是永无止境的，因为一些相似的情况却会产生各不相同的结果。劳资双方在薪金谈判中，不同背景会产生不同的谈判结果；不同的双方也会做出不同的选择。理论家无法有系统地说明为什么一方比另一方获得更多，他们将此结果归因于不明确且难以理解的"谈判力量"差异。

即使是简单的纳什均衡理论也无法对此做出更多解释。假设两人要瓜分1美元。让我们来建构一个博弈，其中每个人都要同时说出自己想要的数目。若两人说出的数目 x 与 y 的总和刚好为1或者小于1，那么他们就能得到各自说出的数目的钱。若两人说出的数目总和超过1，则他们什么也得不到。所以加总为1的任何（x，y）组合就是本博弈的纳什均衡——在给定另一参与人说出的数目的情况下，每个参与人都最好坚持自己说出的数目。[1]

博弈论的进一步发展产生了两条差异很大的研究思路，每一条研究思路都运用了特定的博弈理论推理方法。在第2章中我们区别了合作与非合作的博弈理论：在合作博弈中，参与人共同采取行动；在非合作博弈中，参与人独立采取行动。谈判理论的进一步发展也沿着这两条思路，或运用合作博弈理论，或运用非合作博弈理论。其中一种将谈判视为合作博弈，谈判中的各方共同找出解决方法，或者经由中立的第三方（如仲裁者）来进行解决。另一种则将谈判视为非合作博弈，其中谈判各方独立选择他们的策略，然后达到其均衡。与前面的同时说出自己想要的钱数的简单博弈不同，此处的均衡是不明确的，因此我们得使用更多的方法，进行连续出价和反出价的博弈，使其均衡变得明确。正如第2章所强调的，"合作"与"非合作"是指共同行动和独立行动，并非好的行为和坏的行为，也不是妥协和决裂。非合作谈判的均衡的达成需要很多的妥协。

17.1 纳什的合作解

本节将讨论有关谈判的纳什合作博弈。我们首先利用一个数字实例来介绍这一概念，然后用代数方法进一步阐述一般理论。[2]

17.1.1 数字实例

假设现在有两位硅谷的企业家安迪与比尔。安迪所生产的微晶片每组以900美元卖给电脑制造商。比尔开发的套装软件以100美元零售。二者会面后得知他们的产品能够彼此组合，只要进行细微的调整就可以生产一套硬件与软件的组合系统，每套售价为3 000美元。因此，若他们一起生产，每套将可增加2 000美元的价值，而他们预期每年能够销售数百万套的产品。这样，财富路上唯一的障

碍就是该如何分配这些利润。在每套产品获得的 3 000 美元的收益中，安迪和比尔各该分配到多少钱?

比尔认为，若没有他的软件，安迪的微晶片只是一堆金属与沙粒罢了，所以安迪只应获得 900 美元，而自己则应得到 2 100 美元。安迪反驳说，若没有他的硬件，比尔的软件只是纸上的一堆符号或是磁片上的一些磁性信号罢了，所以比尔只应获得 100 美元，而自己则应获得 2 900 美元。

对他们两人的争论，你可能会建议他们“瓜分差额”。然而这个建议还是不够清楚明确的。比尔或许会提议与安迪平分每套产品的利润。在此方案下，每人可获得 1 000 美元的利润，这表示将每套产品收入额的 1 100 美元分配给比尔，1 900 美元分配给安迪。安迪或许认为所得利润要根据贡献的比例来分配。这样的话，安迪会得到 2 700 美元，而比尔可获得 300 美元。

若他们直接谈判，那么最后协议就会取决于二人的坚持与耐心。若他们试着将争端交由第三方仲裁，则最后决定会依据仲裁者对软件与硬件相对价值的感觉以及双方向仲裁者陈述时所使用的修辞技巧而定。为了明确起见，我们假设仲裁者所决定的分配比例为 4∶1；也就是说，安迪得到 4/5 的剩余，而比尔获得 1/5，或者说安迪获得的剩余为比尔的 4 倍。在这种方案下，总收入如何分配呢？假设安迪总共获得 x，而比尔则总共获得 y，因此，安迪的利润为 $(x-900)$，比尔的利润为 $(y-100)$。根据仲裁者的决定，安迪的利润应该为比尔的 4 倍，所以 $x-900=4(y-100)$，或者 $x=4y+500$。双方可获得的总收入共为 3 000 美元，因此 $x+y=3\ 000$，或 $x=3\ 000-y$。那么 $x=4y+500=3\ 000-y$，或 $5y=2\ 500$，或 $y=500$，从而 $x=2\ 500$。在这种分配方式下，安迪获得的利润为 $2\ 500-900=1\ 600$ 美元，比尔获得的利润为 $500-100=400$ 美元。这就是根据仲裁者的分配方案分配比例为 4∶1 时的结果。

上面我们用代数方程式分析了一个简单的数字实例，这种方法在许多实际应用中非常有用。接下来，我们将继续考察更多的具体例子，分析决定谈判中利润分割比例的因素。

17.1.2 一般理论

假设有两位谈判者 A 和 B，要瓜分总价值 v，当且仅当他们就某种特定分配方案达成共识的时候，才能达到目的。若没有达成协议，双方将独自行动，或是采取其他的行动，从而 A 将得到 a，B 将得到 b。我们称此为他们的最后底线或者在哈佛谈判项目的术语中称为**“谈判协议的最佳替代”**（best alternative to a negotiated agreement，缩写为 BATNA）。[3]通常 a 与 b 皆为零，但一般而言，我们只需假设 $a+b<v$，从而从协议得到的会是正剩余 $(v-a-b)$；若并非如此，那么这个谈判就会没有结果，因为两方面都会运用其外部机会来获得其 BATNA。

考虑以下的规则：每位参与人都会获得其 BATNA 加上一部分的剩余，A 得到的剩余的比例为 h，B 得到的剩余的比例为 k，而 $h+k=1$。以 x 表示最后 A 所得到的数量，以 y 表示 B 最后所得到的数量，我们可得出

$$x=a+h(v-a-b)=a(1-h)+h(v-b),$$
$$x-a=h(v-a-b)$$

以及

$$y=b+k(v-a-b)=b(1-k)+k(v-a),$$
$$y-b=k(v-a-b)$$

我们称这些式子为纳什方程。也可以这样来看这一问题，将剩余（$v-a-b$）以 $h:k$ 的比例分配给谈判者也就是

$$\frac{y-b}{x-a}=\frac{k}{h}$$

或是以斜率截距式表示为

$$y=b+\frac{k}{h}(x-a)=\left(b-\frac{ak}{h}\right)+\frac{k}{h}x$$

为了要将剩余分配完，x 和 y 必须满足 $x+y=v$。关于 x 和 y 的纳什方程式实质就是上面联立方程的解。

图 17—1 是**纳什合作解**（Nash cooperative solution）的几何图形。P 点为最后底线或 BATNA，坐标为（a，b）。所有能将双方剩余以 $h:k$ 比例进行分配的点（x，y）形成一条通过 P 的直线，其斜率为 k/h；这就是我们刚刚提到的直线 $y=b+(k/h)(x-a)$。所有将全部剩余分配完的点（x，y）都在连接点（v，0）和（0，v）的直线上；这条直线就是我们上面推导出的第二个方程 $x+y=v$。纳什均衡解会出现于两线的交点处 Q 点。此点的坐标就是双方达成协议后各自的支付。

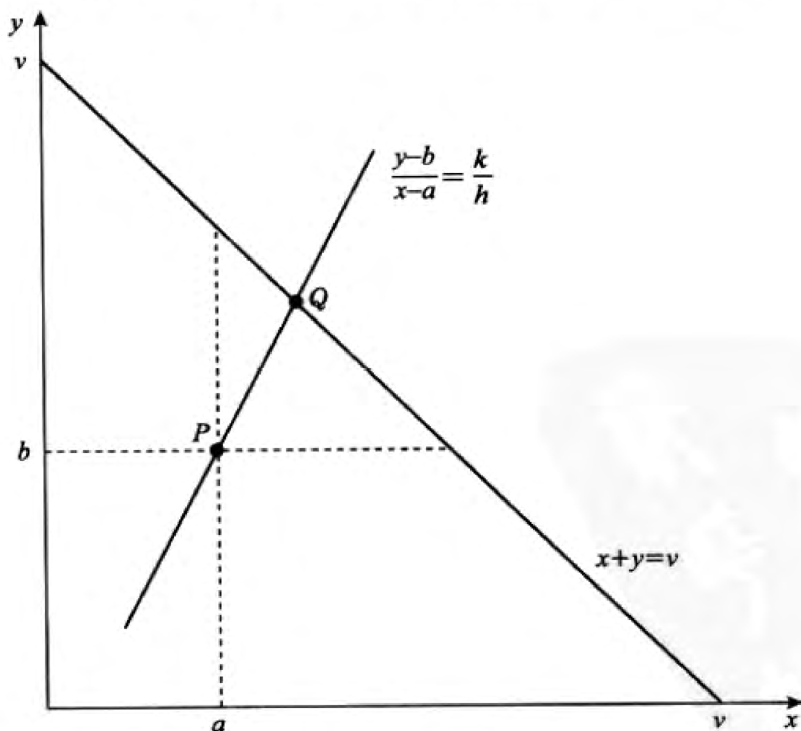


图 17—1　简单例子中的纳什合作解

当然，纳什方程并没有说明此解产生的原因与产生的过程。然而，这种模糊性也正是其优点——它可以被用来简单概括不同的观点和理论。

最简单地说，你可能觉得纳什方程就是谈判过程的简略表示。h 和 k 能够代表双方相对的谈判实力。这个简单的表达式只是权宜之计。完整的理论应该要解释谈判实力是从何而来与为何一方实力大于另一方。我们将在稍后对此加以探讨。我们可以利用此方程作为工具，扼要说明代表谈判力的参数 h 和 k。

纳什使用了不同于目前为止本书所介绍的博弈论方法，因此需要对此做更多说明。到现在为止我们所介绍的博弈中，参与人都各自选择与运用不同的策略。在已知对手的策略的情况下，我们寻找每位参与人获取最佳利益时的均衡点。这样的结果对于一些参与人是非常不利的，甚至对于全体参与人都是不利的，囚徒困境就是最明显的例子。在这种情况下，参与人有机会对共同采取的策略进行协商。但是在我们的框架下，他们无法确保协议能够维持。参与人在达成协议后解散，当轮到该参与人行动时，他仍会采取最符合他自己利益的行动。在面临这样的诱惑时，共同行动的协议往往就会瓦解。的确如此！在第 11 章对重复博弈的讨论中，我们发现关系破裂的潜在威胁可能会维系协议，而在第 9 章中，我们确实考虑了信号的传递。然而，重点在于各自行动，只有相互利益不成为各自行动中利己主义的牺牲品时，双方才可以共同获利。在第 2 章中，我们称此博弈为非合作的，用以强调采取行动的方式，而非共同行动的结果（结果对大家都有利）。因此，此类博弈重要的一点是，各自行动的均衡结果是对大家都有利的结果。

假如共同行动是可能的，那会如何呢？例如，在达成协议后，参与人全部都当场马上采取行动。他们把共同协议或者委托给中立的第三方执行，或是交由仲裁者处理。换句话说，博弈可能是合作的，即共同行动。纳什将谈判视为合作博弈。

共同行动以执行共同协议的团体的思维方式完全不同于一群知道他们正在进行策略性互动但却采取不合作行为的个体。若以均衡的观点讨论后者，则结果的好坏根据他们是否满意其结果而定；至于前者，他们会先思考什么是好结果，然后考虑要如何达成这样的结果。合作博弈的结果是根据理论家认为的合理的普遍原则或特性来定义的。

纳什提出了一整套谈判的原则，并证明了每条原则产生的结果都是独特的。他的原则大致是：（1）若支付呈线性变化，其结果应该不变；（2）其结果应该是**有效的**（efficient）；（3）如果行动选用可能性的降低是因为排除无关因素造成的，那么其结果不会受到影响，因为这些行动无论如何都不会被选用。

第一条原则符合期望效用理论，我们曾在第 7 章附录中扼要讨论过期望效用理论。支付的非线性变化表示参与人对风险改变的态度及其行为上的真正改变；内凹的排列意味着风险规避，外凸的排列则意味着风险偏好，界于这两者中间的线性变化则代表对风险的态度不变。因此，它应该不会影响期望支付的计算，也不会影响期望支付的结果。

第二条原则表示可获得的共有利润应该被获得。在我们所举的简单例子中，A 和 B 要分配总价值 v，亦即意味着 x 和 y 的总和要等于总额 v，而非其他较小

的总额。换句话说，图 17—1 中的解必须落在直线 $x+y=v$ 上。一般而言，谈判的所有可能协议处于所有共同利益都被获取了的子集合中或左下方。此子集合并不需要落在如同 $x+y=v$（或 $y=v-x$）的直线上，它可以落在任何如 $y=f(x)$ 的曲线上。

图 17—2 中，所有在曲线 $y=f(x)$ 上及下方的点组成了完整的可行结果。曲线本身由有效结果所组成；没有任何其他可行结果比曲线 $y=f(x)$ 包含的 x 和 y 的组合更有效，所以在曲线 $y=f(x)$ 上不会存在未被获取的共同利润。因此，我们称曲线 $y=f(x)$ 为谈判问题的**有效边界**（efficient frontier）。例如，若边界线为四分之一的圆弧，$x^2+y^2=v^2$，则 $y=\sqrt{v^2-x^2}$，而可以用 $f(x)=\sqrt{v^2-x^2}$ 来定义函数 $f(x)$。

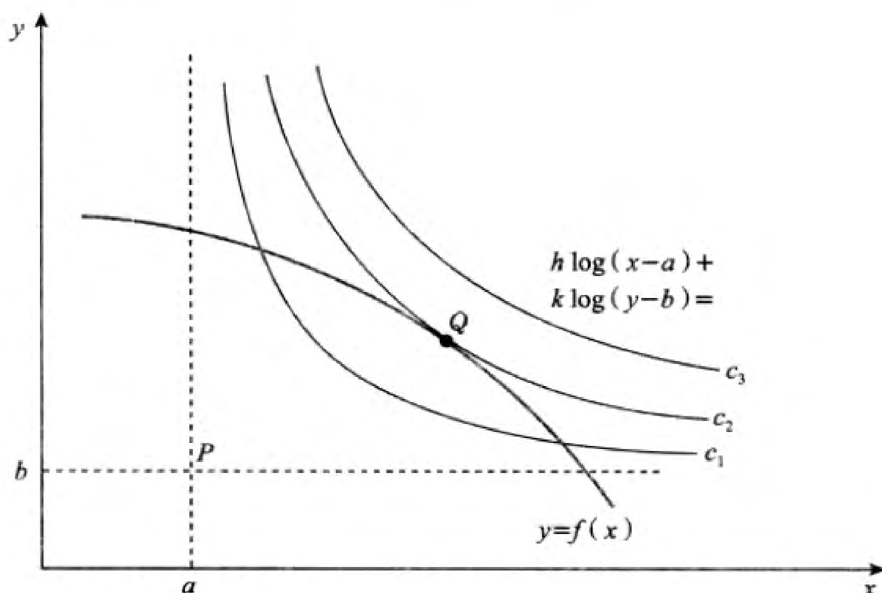


图 17—2　纳什合作解的一般形式

第三条原则似乎也很引人注目。若一位谈判者没有选择图中的任何一点，那结果应该怎样？这个假设与 15.3 节阿罗不可能性定理中“独立于无关候选人”之假设紧密相连，但是我们必须用较高深的方法分析这种联系。

纳什证明符合此三条原则的合作结果可以用数学上的最大化方法来刻画：求解 x 和 y，使之在满足

$$y=f(x)\text{的情况下，最大化}(x-a)^h(y-b)^k$$

这里 x 和 y 皆为结果，a 和 b 是最后底线支付，h 和 k 是总和为 1 的正数，类似于纳什方程中的谈判实力。仅靠纳什的三条原则并不能决定 h 和 k 的值，因此在理论与结果中都有某种程度的弹性。实际上，纳什在问题上加入了第四条原则——两个参与人之间具有对称性。这个额外的假设使得 $h=k=1/2$，并决定了唯一解。我们已经给出了一个具有一般性的方程式，它在随后的博弈论中会广泛地出现。

图 17—2 也对最大化的目标进行了几何阐释。标有 c_1、c_2、c_3 的黑色的曲线是被最大化的函数的无差异曲线；沿着每一条这样的曲线，$(x-a)^h(y-b)^k$ 的

值都是固定的，而如图所示，在 $c_1<c_2<c_3$ 的情况下，$(x-a)^h$ $(y-b)^k$ 会等于 c_1、c_2 或 c_3。整个空间有无限多条类似 c_1、c_2 和 c_3 的曲线，每条都有其固定值，而且愈靠近右上方的曲线，其固定值愈大。

很明显，函数可能的最大值在有效边界与其中一条无差异曲线相切时取得。[4] 通常都是以此切点对纳什合作解进行几何说明。[5]

在图 17—2 中，我们也可利用数学方法推导纳什解。然而此处要表达的目的却比其他方法更重要——至少在博弈理论的学习上是这样的。为了方便，令 $X=x-a$，$Y=y-b$，其中 X 为 A 的剩余，Y 为 B 的剩余。鉴于结果的有效性，有 $X+Y=x+y-a-b=v-a-b$，这就是总剩余，我们以 S 表示，即 $Y=S-X$，且

$$(x-a)^h\ (y-b)^k=X^hY^k=X^h\ (S-X)^k$$

在纳什解中，X 取最大化此函数的值。基本计算告诉我们，若令方程中关于 X 的导数为零，就能找到解。所以我们得到

$$hX^{h-1}(S-X)^k-X^hk\ (S-X)^{k-1}=0。$$

当我们消去公共因子 $X^{h-1}(S-X)^{k-1}$ 后，就得到

$$h(S-X)-kX=0$$
$$hY-kX=0$$
$$kX=hY$$
$$\frac{X}{h}=\frac{Y}{k}$$

最后，以原来的变量 x 和 y 表示方程式，可得 $(x-a)/h=(y-b)/k$，这正是纳什方程。关键之处在于，从纳什的三条原则所推导出的式子就是分配谈判剩余最简单的方法。

用来决定纳什合作解的三条原则或特性是简单而醒目的。但是，由于缺乏完善的监督机制来保证各方按照协议的规定采取行动，因此这些原则可能会无效。参与人若能以策略行动而达到比纳什解更好的结果，他们就会放弃这些原则。若仲裁者强制执行纳什合作解，参与人就可能会干脆拒绝仲裁。因此，若纳什合作解能够解释谈判的非合作博弈，它将会更令人信服。这是有可能的，我们将在 17.5 节中介绍一个重要的特例。

17.2 可变威胁的谈判

在本节中，我们将纳什合作解应用于特定博弈中，也就是序贯博弈的第二阶段。我们在第 17.1 节中曾假设参与人的最后底线（或 BATNA）a 和 b 都是固定的，但是现在假定在谈判博弈的第一阶段中，参与人能够在一定限度内采取策略性行动以操纵其 BATNA。在他们完成此行动后，由 BATNA 产生的纳什合作结

果就会出现在第二阶段中。这种类型的博弈称为**可变威胁谈判**（variable-threat bargaining）。在这种类型的博弈中，参与人应如何操纵 BATNA 才符合其利益？

我们用图 17—3 说明操纵 BATNA 的可能结果。原来已知的最后底线（a 和 b）为博弈最后底线点 P 的坐标；具有这一最后底线的纳什解落在点 Q。若参与人 A 为了将博弈最后底线点移至 P_1 而提高其 BATNA，则从该处产生的纳什解将位于点 Q'，如此将对 A 更为有利（而对 B 更不利）。因此，提高自身的 BATNA 的策略性移动是令人满意的。例如，若你已拥有一个好的工作机会——具有较高的 BATNA，当你再到别的公司面试时，与尚未获得第一份工作时相比，此时你对第二家公司的要求会更高。

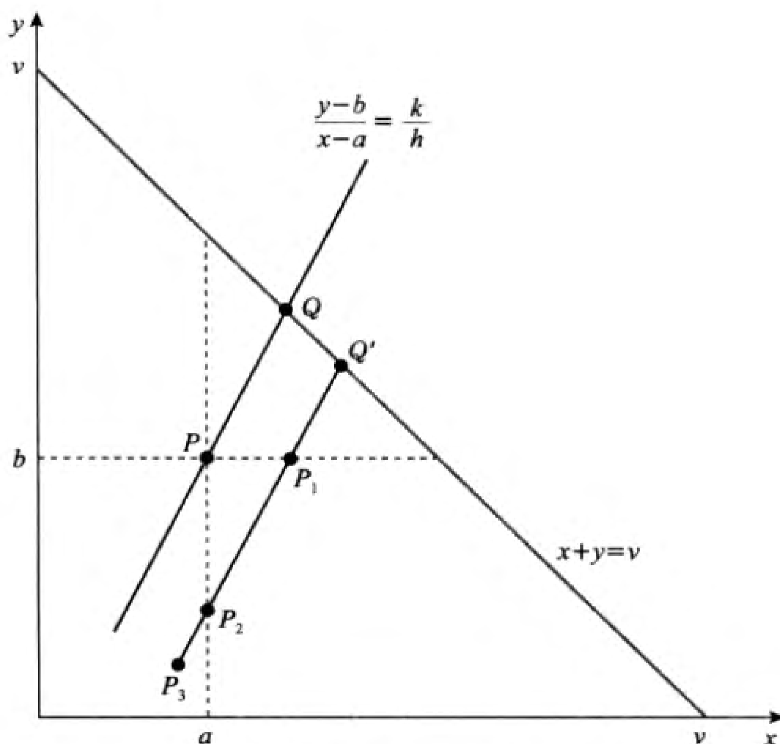


图 17—3　操纵最后底线的谈判博弈

提高 BATNA 可以明显改善最后的结果，然而下一步分析就没有这么明显了。结果表明，如果 A 能够采用降低 B 的 BATNA 的策略性行动，而将博弈最后底线点移至 P_2，则从该处产生的纳什均衡就会落在点 Q'，这和 A 为增加其利益而使最后底线达到 P_1 点时的结果是相同的。因此，与前一种情况相比，这一操作能同样地增加 A 的利益。关于降低对手的 BATNA 的例子，请试想一种情况：你正在工作并想要升职。若你把自己变成公司里必不可少的人，少了你，公司的前景将遭受影响，如此你的升职机会就会大增；如果达不成协议，即公司不升你的职，你就会离开公司——这样可能会让公司利益受损，也就能促使公司同意你的要求。

最后，也是更戏剧化的，若参与人 A 采取行动同时降低两位参与人的 BATNA，从而使得博弈的最后底线点移至 P_3，则又会得到和之前的操作相同的结果。这种移动如同使用威胁："它伤害你会多于它伤害我"。

一般来说，对于A而言，其关键是将博弈的BATNA点移至低于PQ线的某处。BATNA点的位置向右下方移动得越远，A的最后结果会越好。事实上，威胁并不是要造成低支付，而只是以其作为带来更好结果的杠杆。

以这种方式操纵BATNA的可能性要根据其背景而定。这里有个实例——1980年的一场棒球运动员罢工，它呈现出非常复杂的形式。运动员们在春季训练时罢工，然后在4月正常球季开始时恢复打球，而在战争纪念日当天又再度开始罢工。对于劳资双方而言，罢工成本都是巨大的，但是双方成本是不同的。在春季训练期间，球员并没有领薪，而雇主可让球迷参观球员训练而获利。4月及5月正常球季开始后，球员开始薪领，但是由于天气寒冷，所以球季尚未真正到来，观众仍然很少，因此雇主的罢工成本还是不高。球赛观众从战争纪念日后开始增多，所以虽然球员所损失的是相同的薪水，然而，雇主的成本却提高了。因此，这种两阶段罢工是相当聪明的设计，它尽可能地将雇主的BATNA降至比球员的BATNA还低的水平。[6]

还有一件难以理解的事：究竟为何会发生罢工呢？根据此理论，大家都应该知道发生了什么事；解决方式对球员越有利，则越没有罢工的必要。实际中罢工的发生是威胁“走偏”的结果。一些不确定性——非对称信息或边缘政策——应该是罢工发生的原因。

17.3 轮流出价模式一：总价值的损失

我们将研究方向转向更为现实的非合作博弈理论，并思考可能产生均衡的个人实施策略的过程。**一种轮流出价**（alternating offers）模式就是此过程的标准写照。一个参与人A，先提出一个价格，另一个参与人B接受或提出另一个价格，若是后者，则A可以接受此价格或是再提出另一个价格，依此类推。这样，就有一个序贯行动博弈，我们将用逆推法找出其均衡。

用逆推法求解均衡时，我们必须从末端开始往回推。那么哪里是末端呢？为何出价与对手还价的过程最后会停止呢？或者更彻底地问，为何会轮流出价呢？为什么两位谈判者不坚持他们最初的价格，并拒绝改变意见呢？当然，倘若他们无法达成协议，双方将都要付出相当大的代价；但是，对于先让步或做出较大让步者而言，他从协议中获得的好处总是较小。任何人让步的原因，一定都是继续坚持会造成更大的损失。这种损失有两种主要形式：可获得的剩余随着每次出价而**递减**（decay）。这是本节讨论的一种可能性。另一种可能性是，时间具有价值，而**不耐烦**（impatience）在此过程中也扮演了重要角色，故拖得越久协议越没有价值。我们将在17.5节中考察此种可能性。

考虑下面剩余不断递减的例子。一位球迷抵达了一场职业球赛赛场门口，但是他没有球票。他最多愿意付25美元观看每节比赛。他发现球场外有个票贩在倒卖球票。票贩向他提出一个价格，球迷若不愿意支付此价格就会转身到附近的酒吧里，在大屏幕电视上观看第一节球赛。等这节结束时，他走出酒吧，发现票

贩还在那里，于是球迷就出了另一个价格。若票贩不同意，球迷就会回到酒吧里。第二节结束时，他会再度走出酒吧，而票贩会再提出另一个价格。倘若球迷还是无法接受此价格，他会再次回到酒吧，并在第三节结束时再向票贩出价。当然，现场观看球赛的价值会随着所剩节数减少而降低。[7]

逆推分析可让我们预测这种轮流出价谈判的结果。在第三节结束时，球迷知道此时若他不买球票，票贩的球票将会变成一张废纸。所以即使球迷所提出的价格非常低，票贩还是会把票卖给他，因为聊胜于无。因此在球迷最后出价时，他几乎可以免费获得球票。反推第二节结束时，票贩握有出价的主动权。但是他必须考虑到后来的情况，并认清一个事实——他无法从球迷那儿获得剩余两节球赛的完全价值。若票贩的出价高于 25 美元（球迷对第三节球赛的评价），球迷就会拒绝票贩的价格，这是因为球迷知道他将可以用几乎为 0 的价格观看第四节的球赛；所以，票贩的最高出价只能是 25 美元。

现在考虑第一节球赛结束时的情况。球迷知道若他现在不买下球票，之后票贩的期望价格仅为 25 美元，所以球迷只要出价 25 美元就可获得球票。最后，在比赛正式开始前，票贩会考虑到后来所有的状况并出价 50 美元。此 50 美元代表的是，球迷对第一节球赛 25 美元的评价与对剩余三节球赛 25 美元的评价。因此，两者会立刻达成协议，以 50 美元成交，这个价格是完全由逆推法得出的。[8]

我们可以将这个例子简单地抽象成两位谈判者 A 和 B 间的一般情况。假设 A 首先出价要瓜分价值为 v 的总剩余。若 B 拒绝此价格，那么可获得的剩余就减少 x_1 降至 $(v-x_1)$；然后由 B 出价。若 A 拒绝 B 的出价，那么可获得的剩余就减少 x_2 降至 $(v-x_1-x_2)$；再由 A 出价。然后彼此再继续出价，一直到最后，我们假设 10 个回合后博弈就结束了，即 $v-x_1-x_2\cdots-x_{10}=0$。如同一般序贯行动博弈，我们在博弈结束处开始进行分析。

博弈进行到只剩下 x_{10} 的剩余时，B 可以提出最后价格，将“几乎所有的”剩余留给自己，只分极少部分给 A。此时由于除此之外别无选择，而 A 不得不接受这种出价。为避免斤斤计较几分几厘，让我们称这个结果为“B 获得 x_{10}，A 获得 0”。我们将对前面的回合进行相同的分析。

知道了博弈的第 10 个回合里发生了什么事，我们再来考察第 9 个回合。在第 9 回合的博弈中，是由 A 出价，而此时剩余为 (x_9+x_{10})。A 知道他至少要对 B 出价 x_{10}，不然 B 会拒绝其价格，从而进入第 10 回合。当然，A 不愿意给 B 更多的剩余。所以在第 9 个回合中，A 的出价会使他本身能保留 x_9，而留 x_{10} 给 B。

所以在更早的一个回合中，当剩余为 $x_8+x_9+x_{10}$ 时，B 的出价会使 A 得到 x_9，让自己能获得 (x_8+x_{10})。再逆推回去，第 1 个回合中，A 的出价会使他自己能获得 $(x_1+x_3+x_5+x_7+x_9)$ 而留下 $(x_2+x_4+x_6+x_8+x_{10})$ 给 B。这样的出价是会被接受的。

你可以用简单的技巧记住这个例子。假设有这样一种结果，即所有的出价都被拒绝（当然这不可能真的发生），然后，加总所有被某一参与人拒绝的价格，所得到的总数就是在真正均衡时，另一参与人会得到的剩余。如上例所示，当 B 拒绝了 A 的第一次出价，可获得的剩余就会减少 x_1，而 x_1 就会成为博弈均衡时

A 可获得的剩余的一部分。

若每位参与人都有一个正的 BATNA，则应将其纳入考虑之中，分析也因而必须加以修正。在最后一个回合中，B 向 A 的出价至少必须等于其 BATNA a，若 x_{10} 比 a 大，那么这会让 B 获得（$x_{10}-a$）；若 x_{10} 比 a 小，那么博弈就会在达到这一回合之前结束。在第 9 个回合中，A 向 B 提出的价格必须是 B 在第 10 个回合中能得到的（$x_{10}-a$）和不能达成协议时 B 所能获得的 BNTNA b 这两个值中的较大的一个。可以利用这种方法一直逆推回第一个回合；我们将剩下的工作留给你，请你完成其余的逆推。

我们已经找到了轮流出价谈判博弈的逆推均衡，而在推导结果的过程中，我们也说明了存在于均衡后面的策略——其实是完整而有条件的行动，也就是若博弈在之后的回合达成协议，每位参与人会怎样做。实际上，真正的协议在一开始出价时就会立刻达成。所以之后的回合根本无法达成协议，它们是均衡点外的节点与路径。和逆推法一样，对理性行为者在其将会达到的节点上的行动的预想催生了最初的行为。

另一个重点是剩余的逐渐减少（在数个回合的出价后），这比突然减少（只有一个回合的谈判）公平。在后者中，若 B 拒绝 A 的第一次出价，协议就无法达成；而在逆推均衡中，A 会企图将（几乎）所有的剩余都留给自己，而向 B 下达"最后通牒"，要 B 接受极少或甚至为 0 的剩余。然而，连续回合的出价让 B 有能力拒绝不公平的第一次出价。

17.4 实验证明

此类谈判程序很简单，实验室或课堂实验通常通过构造总剩余递减的条件来观察实验主体的真正行为。17.3 节中，我们在讨论逆推法的有效性时就已扼要提及了相关内容，而现在我们要在谈判的背景下，更详细地考察这些例子。[9]

最简单的谈判实验是**最后通牒博弈**（ultimatum game），这是仅有一个回合的博弈：由参与人 A 出价，若参与人 B 不接受其价格，谈判就结束了，结果双方都没有获得任何东西。这个博弈的一般情况如下。将一群参与人置于同一间教室里，或是置于同一网络中，两两被配对；然后指定每一对中的一人为出价者（提出价格的人，或定价格的卖方），另一人被指定为选择者（接受价格或拒绝出价的人，或决定是否以该价格购买的顾客）。每一对的固定剩余为 1 美元或是其他数目，而他们要瓜分剩余。

逆推法表明，A 向 B 的出价应为最小的货币单位，比如说 1 美分，而 B 应该接受此价格。然而，真正结果是完全不同的。在本例中，由于将实验对象聚集在同一间教室里，并随机指定出价者，因此最常见的出价为以 50 : 50 的比例平分剩余，很少会有比 75 : 25 的比例更糟的状况（出价者得 75%，而选择者得 25%）。如果这种状况真的发生，选择者通常是不会接受的。

对这项发现有两种解释。参与人可能无法进行逆推法必需的计算或没有进行

此计算，或者他们的支付包含了此回合博弈以外的东西。当然，最后通牒博弈的计算很简单，每个人都能计算，因为这些实验的主角大部分都是大学生。一种较可能的解释即我们曾在 3.6 节和 5.3 节里阐述的理论。这一理论假设支付仅为从该回合谈判中获得的收益，过于简单。

参与人的支付还可能包含其他东西。他们可能由于自尊而不愿意接受很不公平的分配方案。即使出价者 A 并没有将这个因素纳入对自己的支付的考虑之中，但若他觉得 B 可能会被此所影响，那么 A 最好还是将价格提高到 B 能够接受的范围内。向 B 提出较低的出价会抵消 A 的高支付，因为 B 会拒绝接受太不公平的出价从而使得 A 什么也得不到。

第二种可能的解释是，由于实验中的参与人都在同一间教室，从而无法保证匿名配对。若参与人是来自同一群体的大学生或是同一个社区的居民，则他们都会有博弈以外的关系，那么他们可能会考虑这些关系。因此，出价者会担心他们在博弈中的出价过于不公平将会损害他们在博弈以外的关系。所以，比起简单的博弈来，此时的出价者就会大方得多。倘若这就是解释，那么可以肯定更多的匿名者将使得出价更不公平。实验的确证实了这一点。

最后，人们在成长或受教育的过程中会被灌输公平的观念。公平的观念对于整个社会而言具有演化价值，也可能因此而变成社会规范。无论其根源如何，也无论其是否会被拒绝，出价者的出价都变得相对大方。作者之一（斯克丝）曾在教室内进行过几种最后通牒博弈的实验。在彼此认识的学生间进行瓜分谈判时，谈判结果都明显较公平。此外，有些学生用特定文化背景解释了异于理论预测的行为。

我们已经进行了许多基本博弈的变化形式的实验以区别不同解释。博弈外的关系可用严格的匿名程序来规范。这样做虽然会影响结果，但却不会产生像纯粹自私的逆推理论所预测的那样极端的结果。我们特别挑出“担心被拒绝”及“根深蒂固的公平观念”的情况来加以讨论。

考虑一种变异的专制博弈，这种博弈消除了担心被拒绝的顾虑。参与人再一次地被配对，我们指定 A 来分配，而另一位参与人 B 则为被动的接受者。现在的分配会更明显地不公平，但大多数的 A 分配给自己的仅仅为 70%。这样的结果令人联想到根深蒂固的公平观念。

然而，这种观念有其局限性。在一些实验中，公平观念只会产生于随机指定出价者与选择者的时候。现在稍加变化，比如规定参与人参加一个简单测验，表现较佳者会成为出价者。这个测验使参与人产生“出价者”的身份是竞争而得的感觉，而实验结果也的确会产生更不公平的分配。当专制博弈的实验是在赢的前提与严格匿名条件下进行时，大部分的 A 都保留所有的剩余，只有一些（大概 5%）A 仍然会以 50∶50 的比例进行分配。

作者之一迪克西特曾经做过一项实验，他将 20 名学生聚集于电脑教室里，并将他们随机匿名地配对，每对学生都就如何分配 100 分进行谈判。他并不指定出价者与选择者的角色；任何一方都可以出价，也都可以接受对方的分配。随时都能出价或改变价格。每一对学生能利用电脑即时地与其他同组对手交换信息。谈判随机于 3～5 分钟后结束；若双方没有在这一时间内达成协议，则两

人都获得 0 分。他以不同的随机配对进行了 10 个回合的这样的博弈，这样参与人不会因为反复博弈而有合作的机会。在教室实验中，学生们都会有课堂以外的私人关系；但是在这种实验里，虽然并非匿名配对，学生们却无法得知或猜到他们在每个回合中将扮演的角色。在整个博弈中，每位学生的得分就是 10 回合所得分数的总和。因为这一数值太高了点，所以将总和的 5%列为本科目的学科成绩。

最高的总分为 515。表现最好的是很快达成以 50∶50 的比例分配的协议的参与人，而那些企图获得不公平的分数的学生、拒绝分配差异在 10 分之内的学生，或者考虑太久而逾时者，都表现得很差。[10]就一方的支付而言，似乎中庸与公平者的表现较佳。

17.5 轮流出价模式二：不耐烦

本节中，我们将讨论另一种因延迟达成协议而产生的成本问题。假设用于进行分配的利益本身的货币价值并不会递减，但是参与人有“货币的时间价值”，他们偏好较早一点达成协议。与 17.3 节描述的一样，他们轮流出价，但是现在他们都偏好早一点得到钱甚于晚得到钱。具体而言，就是谈判双方都相信，现在的 95 美分好过晚一点的 1 美元。

偏好马上获得好于晚一点获得的行为，就是不耐烦，即现在的重要性比未来大。我们在 11.2 节讨论过此概念，并提及了两个原因。第一，假设参与人 A 投资 1 美元，收益率为 r，因此下一期（明天、下礼拜，或明年，任何时间够长的一期）的所得总额为 $(1+r)$。第二，风险的存在会让博弈可能在本期与下一期出价之间结束（如同之前的教室实验中，博弈会随机于 3～5 分钟内结束）。若 p 为博弈继续进行的可能性，则现在对下一回合能够获得的 1 美元的期望值为 p。

假设谈判双方的 BATNA 皆为零。让我们由其中一个参与人开始这个博弈，令 A 先出价分配 1 美元。若另一参与人 B 拒绝 A 的出价，则 B 有机会在下一回合里出价。出价时，双方的情况是完全相同的，因为要分配的总额一直都是 1 美元。因此无论是 A 还是 B，在均衡时，出价者获得的都是一样的（称为 x）。我们可以利用逆推法找出能解出 x 的方程式。

假设轮流出价的程序由 A 开始。他知道当下一回合由 B 出价时，B 将可以得到 x。因此，A 对 B 的出价，其大小至少要等于下一回合的 x，所以 A 现在至少要给 B $0.95x$。（请记住，对于 B 而言，在本回合获得的 95 美分等于下一回合所得的 1 美元，因此现在的 $0.95x$ 等于下一回合的 x。）当然，A 不会提供更多给 B 以引诱他接受。所以 A 对 B 的出价为 $0.95x$，A 本身留下 $(1-0.95x)$，但是当 A 出价时，其所得刚好就是我们前面所提到的 x。因此 $x=1-0.95x$，或 $(1+0.95)x=1$，或 $x=1/1.95\approx 0.513$。

计算时，请你注意两件事。第一，虽然此过程允许无限次的出价与还价，但均衡状态下 A 第一次出价时，B 就会接受该价格，博弈因此结束。因为时间是有

价值的，所以这是一个有效结果。延迟的成本将根据 A 要给 B 多少钱以引诱 B 接受其出价而定。这会影响到 A 的逆推过程。第二，先出价者能够获得一半以上的支付，即为 0.513，而不是 0.487。因此，任一参与人自己率先出价时的所得都高于对方率先出价时的所得。但是，在最后通牒博弈中没有还价的机会，因此，与之相比，此时这一优势小得多。

现在假设两个参与人并非一样有耐性（或不耐烦）。B 还是认为下一回合的 1 美元和现在的 95 美分具有同等价值，但是 A 却将下一回合的 1 美元当做现在的 90 美分钱。因此 A 愿意接受更少的钱以求得快点达成协议；换言之，A 比较没有耐性。这种耐性上的差异可以解释谈判的均衡支付的不平等。为找出此例的均衡点，我们以 x 表示 A 首先出价时所获得的总额，而用 y 表示 B 首先出价时所获得的总额。

A 知道他对 B 的出价至少要为 $0.95y$，不然 B 就会拒绝接受，因为他知道当轮到 B 出价时，他能够获得 y。因此，A 所得的 x 必须为 $1-0.95y$；故 $x=1-0.95y$。同样地，当由 B 开始出价时，他知道他对 A 的出价至少要为 $0.90x$，所以 $y=1-0.90x$，经由这两个方程式，可以解出 x 和 y：

$$x=1-0.95(1-0.9x)$$
$$[1-0.95(0.9)]x=1-0.95$$
$$0.145x=0.05$$
$$x=0.345$$

和

$$y=1-0.9(1-0.95y)$$
$$[1-0.9(0.95)]y=1-0.9$$
$$0.145y=0.10$$
$$y=0.690$$

请注意，x 和 y 加起来并不等于 1，因为这两个值都是每个参与人率先出价时的支付。因此，当 A 率先出价时，A 得到 0.345，而 B 得到 0.655；当 B 率先出价时，B 得到 0.69，而 A 得到 0.31。也就是说，每个参与人在他自己率先出价时的所得会多于在对手率先出价时的所得，但是差异很小。

参与人不耐烦程度不相同时的结果与参与人不耐烦程度相同时的结果大相径庭。当不耐烦程度不相同时，较没有耐性的 A 所得会较少，即使是他率先出价。所以一个愿意接受较少利益以快点获取这一利益者，最后所得一定较少，并且差异是很大的。上例中 A 对 B 的比例接近 1∶2。

如前所示，我们可根据这些例子推导出更一般化的代数式。假定 A 将立即得到的 $1/(1+r)$ 美元和之后得到的 1 美元视为具有相等价值，在计算中我们以 a 替代 $1/(1+r)$。同样地，B 将立即得到的 1 美元的价值视为等于之后得到的 $(1+s)$ 美元，我们以 b 替代 $1/(1+s)$ 美元。若 r 比较大（或若 a 比较小），那么 A 就是相当不耐烦的。类似地，如果 s 较大（或若 b 较小），B 就是不耐烦的。

我们现在来看看轮流出价时的情况。两位谈判者要分配 1 美元，两者的 BATNA 皆为零。（若你了解此例，你就能够使用更一般的例子。）本例的逆推均衡是什么？

我们可以扩展稍早使用的简单论点以找出均衡时的支付。假设当 A 率先出价时，其逆推的均衡支付为 x；当 B 率先出价时，其逆推的均衡支付为 y。我们若能找到关于 x 和 y 的方程式，就可以解方程式，求出均衡支付。[11]

当 A 率先出价时，他知道给 B 的价格必须等于下一回合的 y。这一价格为

$by=y/(1+s)$。在出价后，A 可以保留 $x=1-by$。

同样地，当 B 率先出价时，他知道给 A 的价格必须等于下一回合的 x，即 ax。因此 $y=1-ax$。现在不难解这两个方程。所以我们得到 $x=1-b(1-ax)$，或 $(1-ab)\ x=1-b$。将其以 r 和 s 表示，有

$$x=\frac{1-b}{1-ab}=\frac{s+rs}{r+s+rs}$$

同样地，$y=1-a/(1-by)$，或 $(1-ab)y=1-a$，则有

$$y=\frac{1-a}{1-ab}=\frac{r+rs}{r+s+rs}$$

虽然求解需要一些技巧，但是这与之前使用的步骤是一样的，我们随后将会用另一种推导方式得到同样的答案。先让我们考察这个答案的一些特性。

请注意，在先前不耐烦程度不相等的例子中，x 和 y 的总和大于 1：

$$x+y=\frac{r+s+2rs}{r+s+rs}>1$$

请记住 x 是 A 有权率先出价时所得到的支付，而 y 则是当 B 有权率先出价时所得到的支付，当 A 率先出价时，B 得到比 y 小的 $(1-x)$；这显示出 A 率先出价时的优势。同样地，当 B 率先出价时，B 得到 y 且 A 得到比 x 小的 $(1-y)$。

然而，通常 r 和 s 都很小，当出价可以在很短的时间内（例如一周、一天、或一小时）完成时，参与人获得的利润就会非常小，而且博弈恰好在下一个回合结束的可能性也是非常小的。例如，若 r 是 1%（或 0.001），而 s 是 2%（或 0.002），则由方程可得 $x=0.668$ 与 $y=0.337$，所以第一次出价的好处仅为 0.005。（A 在自己率先出价时获得 0.668，而在 B 率先出价时获得 $1-0.337=0.663$；所以差异仅为 0.005。）更正式地说，与 1 相比，r 和 s 值很小时，rs 也会非常小，因此，我们可以将 rs 忽略掉，从而得出一个近似解，这一解与谁率先出价无关：

$$x=\frac{s}{r+s} \qquad \text{和} \qquad y=\frac{r}{r+s}$$

现在 $x+y$ 等于 1。

最重要的是，近似解里的 x 和 y 就是双方的剩余占有率，而且 $y/x=r/s$，也就是说，参与人的剩余占有率，与以 r 和 s 所衡量的不耐烦程度呈反比例关系。若 B 的不耐烦程度为 A 的 2 倍，则 A 所得到的剩余会是 B 的 2 倍，所以占有率分别为 1/3 和 2/3。因此，耐性是谈判中的重要优势。正式的分析支持如下观点：若你很没有耐性，另一位参与人就会非常干脆地提出一个较低的出价，并且知道你会接受它。

不耐烦令美国政府机构和外交官在与他国的许多谈判中吃了大亏。美国政治程序强调速度。媒体、利益团体与政敌都会要求得到结果，它们往往都习惯于因延迟而批评行政当局或外交官。在这种压力下，谈判者都想快点得到结果。从美国长期的经验来看，这样得到的结果都是未经深思熟虑的，他国的让步通常都是

有漏洞的，因此承诺的置信度也就不高。美国行政当局却把这种交易当做胜利，但通常在数年以后就不了了之。一个相似而更生活化的例子是：保险公司常给予遭受重大损失的保险人很低的补偿额，因为它们知道客户都急于得到结果，因此都很没有耐性。

就概念上而言，方程 $y/x=r/s$ 将非合作博弈的谈判方式与 17.1 节的纳什合作方法联系在了一起。关于可获得的剩余的占有率的方程，在 BATNA 等于零的情况下为 $y/x=k/h$。在非合作方法中，双方的占有率与他们的谈判实力成正比，然而实力被假设是由外在方式所决定的。现在，以参与人的基本特性来解释谈判实力时，k 和 h 与双方不耐烦程度 r 和 s 成反比。换句话说，我们可以参与人的特性（如不耐烦）正确解释纳什合作解中抽象地表征谈判实力的参数，纳什合作解可以被看成非合作轮流出价逆推均衡的一种解释，且这种解释是较令人满意的。

最后，请你注意协议是即时达成的——最初一次的出价将会被接受。整个逆推分析必须符合第一次出价的特性，即率先出价者要知道，对手一定会拒绝太低的出价。

我们利用另一种方式来推导稍早获得的出价均衡方程式，并以此结束本节的阐述。考虑一个有 100 个回合的讨价还价过程，其中 A 为先出价者，而 B 为后出价者。从第 100 个回合开始进行逆推分析，此时 B 会保留全部的 1 美元。因此，在第 99 回合中，A 对 B 的出价，其价值要等于在第 100 个回合中的 1 美元，即 b，而 A 将保留 $(1-b)$。依此逆推这一过程：

在第 98 个回合，B 对 A 出价为 $a(1-b)$，而 B 保留 $1-a(1-b)=1-a+ab$。

在第 97 个回合，A 对 B 出价为 $b(1-a+ab)$，而 A 保留 $1-b(1-a+ab)=1-b+ab-ab^2$。

在第 96 个回合，B 对 A 出价为 $a(1-b+ab-ab^2)$，而 B 保留 $1-a+ab-a^2b+a^2b^2$。

在第 95 个回合，A 对 B 出价为 $b(1-a+ab-a^2b+a^2b^2)$，而 A 保留 $1-b+ab-ab^2+a^2b^2-a^2b^3$。

以此方式逐步逆推，我们将看到在第 1 个回合时，A 会保留

$$\begin{aligned}&1-b+ab-ab^2+a^2b^2-a^2b^3+\cdots+a^{49}b^{49}-a^{49}b^{50}\\&\quad=(1-b)[1+ab+(ab)^2+\cdots+(ab)^{49}]\end{aligned}$$

很明显，还可以进行更多回合的博弈，并会得到愈来愈多的包含 ab 的项。为了找出 A 作为这场无限轮流出价博弈中的率先出价者时的支付，我们必须先找出此无穷几何级数的极限。第 11 章附录说明了如何求级数的和。根据此处我们所得到的方程，可得到结果如下：

$$(1-b)[1+ab+(ab)^2+\cdots+(ab)^{49}+\cdots]=\frac{1-b}{1-ab}$$

这刚好就是我们稍早所得到的关于 x 的解答。同样道理，你也可以得到当 B 为率先出价者时的支付。进行这种推演时，请同时加强你的理解和计算技巧。

17.6 在谈判中操纵信息

我们已知道谈判结果主要依赖于双方不同的特性，特别是他们的 BATNA 和不耐烦的程度。我们继续假设参与人都了解对方与自己的特性。并且，我们假设每个参与人都知道对方知道自己知道对方的状况，也就是说，特性是大家的共同知识。实际上，我们常在不了解对方的 BATNA 或其不耐烦程度的情况下参与谈判，有时候甚至我们也无法准确得知自身的 BATNA。

我们在第 9 章知道，具有不确定性或信息不对称的博弈与用于操纵信息的信号传递和甄别策略有关。谈判中恰恰充满了这样的策略。拥有颇高的 BATNA 或有耐性的参与人会传递信号告诉大家其特性，然而不具备这些良好特性的参与人也会模仿这样的行为，因此其他人便会产生怀疑，从而严格检验该信号的置信度。因此，各方都会试着利用策略与甄别机制来使他方采取行动，借此识别其真实特性。

在本节中，我们考察卖方和买方在房地产市场中运用的策略。大部分美国人一辈子中都会数次积极投入房地产市场，因此许多人都是经验丰富的专业房地产代理人或经纪人。房地产是美国仍然在期望价格与接受价格之间进行讨价还价的少数市场之一。因此，从中可以获得丰富的策略经验。我们可以将这些经验用于其他例子中，并以博弈理论的观点来进行说明。[12]

当你考虑在一个陌生的地区购买房屋时，你不太可能知道你中意的房屋的一般价格范围。你应该先找出价格范围，然后才能决定你的 BATNA。这并不表示你要阅读报纸上的广告或查看房地产商的清单，因为上面有的仅是要价。地方性报纸或网站则会列出最近交易的实际价格。你应该查询针对相同房屋的要价的还价，以了解市场状况以及可能的谈判范围。

下一步就是要找出（甄别）另一方的 BATNA 与不耐烦程度。若你是买家，你要知道此房屋被出售的理由，以及该房屋在市场上停留了多久。若其为空屋，原因是什么？此房屋被闲置了多久？若房主是由于离婚或搬家至其他地区而空下此房，并正在偿还另一栋房屋的巨额贷款，则他的 BATNA 可能较低或他相当不耐烦。

你应该找出关于另一方偏好的其他特性，即使这些偏好是非理性的。例如，有些人会把低于要价太多的出价视为侮辱，所以无论之后该买方提出任何价格，他都不可能买到该房屋。这类行为因地区和时间变化而改变。花点时间找出具有普遍性的做法是值得的。

更重要的一点，与其他方式相比，接受出价更能够准确暴露参与人的真正意愿，因此这可让对手利用。我们有个聪明的朋友，他是位博弈理论家，刚好经历过一件事。他为一盏立灯而与卖方进行谈判。开始时卖方要价 100 美元，当谈判进行到我们的朋友的出价 60 美元时，卖方说“卖了”，这时他想：“这家伙愿意以 60 美元把立灯卖给我，那么他的底线一定会更低。让我来找出他的底线究竟

在哪里。”所以他说：“那 55，怎么样?”卖方变得很生气，且拒绝以任何价格卖出立灯，接着他就要求我们的朋友离开这家店，并且永远都不要再回去。

卖方的行为遵从这样的准则：在谈判中，任何一方出尔反尔都会造成严重的信用危机。这成为社会中谈判博弈的准则是有其合理性的。若一方的公开出价无法让另一方恪守承诺，也就是谈判无法在不用担心被试探的情况下（如我们的朋友）进行，那么每位谈判者就会在等待对方接受价格时泄露其真实的底线，然后整个谈判过程就会突然停止。因此，这种行为要被禁止，而且要像本例中的卖方一样主动地制止这种不讲诚信的行为，并使这样的制止行为成为一种社会准则，从而形成良好的社会风尚，这才是实现社会目标的最佳方式。

当然，出价可能仅在指定或有限的时间内公开，这项规定可以是出价本身的一部分。招聘通常会有接受简历的截止时间；商店的拍卖也是在某一特定时间之内进行的。然而当出价是价格与时间的组合时，任何方面上出尔反尔都会引起愤怒。例如，顾客在促销期间到商店里消费，却找不到宣传单上的特价物品，他们一定会很生气。因此，商家必须确保下次以一般价格进货时，这些顾客还是能以特价购买该商品；就算这种安排会引起不便，或冒商誉受损的风险，商家还是要这样做。或者这家商店事先在促销宣传单上注明“商品限量，售完为止”；即使如此，顾客发现商店已卖完特价商品时，还是会很沮丧。

接下来要介绍的是用于一对一谈判的策略，类似于在房地产市场上，你可以传递出表示你有高的 BATNA 或有耐性的信号。表现有耐性的最好方式就是显得有耐性，不要太快回应对手的出价。“让卖方觉得已失去你这位买主。”这是一个可信的信号，因为不是真正有耐心的人会觉得模仿这种从容方式的成本太高了。同样地，你可以就这么走开，来表示你有高的 BATNA，这是在其他国家的市场与美国的跳蚤市场中常用的策略。

即使你的 BATNA 很低，你仍可以不接受某个水平以下的价格。这样的限制就像是具有高 BATNA，因为对方无法让你接受任何太低的出价。在房屋谈判中，你可以借口小气的父母无法继续支付预期款，或你的配偶不喜欢这间房屋而不让你多出价，用以表示你无法再做让步了。卖方也能够采用相同的策略。薪资谈判之前一个程序是授权谈判。工人召开会议，以通过授权决议案的方式批准工会领袖代表劳方与资方进行协商，而决议案规定工会领袖不能接受低于某个底线的提议。所以在与资方协商时，工会领袖可以说劳方立场是一致的，没有时间回去问劳方成员他们是否接受任何更低的提议。

大部分策略都伴随着某些风险。当你正以等待表示你有耐心时，卖方可能会找到另一位有意愿的买主。当雇主与工会都在等待对方让步时，紧张情绪会上升而使罢工爆发，这对于双方而言都是成本惨重但却又无法事先预防的情况。换句话说，许多操纵信息的策略都是边缘政策的例子。第 14 章中谈到这种博弈结果不利于双方，在谈判中也是一样。以谈判破裂或罢工作为威胁，都是策略性的行为，都是企图较快达成协议，或是促使参与人采取行动达成较佳的交易，然而，真正的谈判破裂或罢工都是这种威胁“出错”的结果。当然，做出威胁的参与人（即实施边缘政策的一方）在决定是否要继续走这条路且走多远时，必须评估风险与可能的收益。

17.7 多方参与和多项议题的谈判

迄今为止，我们讨论的博弈仅限于典型的情况，也就是两方对剩余进行分配的谈判。但是实际生活中的谈判，大多同时有多方参与，或是涉及多项议题。虽然博弈会变得复杂，然而参与人或议题的增加往往使得达成协议更为容易。在本节中，我们将扼要地讨论此主题。[13]

17.7.1 多项议题的谈判

从某种意义上来说，我们已经思考过涉及多项议题的谈判。买方与卖方对价格的谈判永远会涉及两件事：(1) 买卖双方交易的物品；(2) 钱。当买方对物品的估价大于卖方时，那么潜在的共同利益会提升，也就是买方愿意用来交换此物品的钱数高于卖方愿意接受的数额。双方可以获得较佳的协议结果。

相同的原则可以更广泛地使用。国际贸易就是典型的例子。请试想有两个国家：F 国和 I 国。若 F 国可以用 1 条面包交换 2 瓶酒，而 I 国可以用 1 瓶酒交换 1 条面包，它们就可以"无中生有"地"制造"一些物品。例如，假设 F 国少生产 100 条面包，它就可以多生产 200 瓶酒；而 I 国少生产 150 瓶的酒，它就可以多生产 150 条面包。与两国原本的产量相比较，此资源效用的转换创造了额外的 50 条面包与 50 瓶酒。若两国能够就如何分配达成协议，那么这些面包与酒就是它们能够创造的"剩余"。例如，假设 F 国给 I 国 175 瓶酒，而得到 125 条面包，则两国各会比以前多出 25 条面包与 25 瓶酒。存在一个对应于各种交换的可能范围。在其中一种极端情况下，F 国会放弃它生产的全部 200 瓶酒以交换 I 国的 101 条面包，这种情况下，I 国就几乎获得了交易中全部的利润。在另一种极端情况下，F 国可能只会放弃 151 瓶酒以交换 I 国的 150 条面包，这时 F 国就能够获得几乎所有利润。[14]在此两种情况之间，两国可以对贸易利润的分配方案进行谈判。

一般原则现在就变得清楚了。当谈判桌上同时有两个或两个以上的议题，而双方也愿意以不同比例来进行交换自身较为富余的商品时，互惠的交易就会达成。只要交易比例界于双方意愿比例之间，双方就能够共同获得利益。剩余的分配是根据交易比例而定的。交易比例愈靠近哪一方的意愿比例，哪一方赚钱愈少。

现在你可以明白增加议题是如何同时增加互利交易成交的可能性的了。议题愈多，你愈有可能发现双方在不同议题上的比例差异，也愈容易找到共同获利的可能性。在房屋的例子中，对于要搬家的卖方而言，许多房屋设备及家具或许都没有什么用处了，但是对于买方而言，这些东西可能是他们所需要的。所以，若卖方不愿降低房屋售价，为了成交，卖方或许可以让其售价包含这些物品。

然而，增加议题并非只有好处。若你对某物估价很高，你可能不愿将其置于谈判桌上；你会担心对方因为知道你想要维护此物品的价值而使你让步太多。最糟糕的状况就是，一方利用新增的谈判议题来进行威胁，以降低另一方的 BATNA。例如，正在从事外交谈判的国家，可能会因经济封锁而遭受重大影响，因此其必然偏好将政治议题与经济议题分开讨论。

17.7.2 多方参与的谈判（多边谈判）

多方同时参与的谈判比较容易达成协议，这是因为各方可以找到让步的循环，而不用找出刚好配对的交易。还是以国际贸易为例。假定美国小麦的生产能力较强，但是汽车的生产能力较弱；日本对于生产汽车很有效率，但是却没有石油；沙特阿拉伯拥有大量石油，但是却没有小麦。若以两两配对谈判，它们想达成协议必定很困难，但是若三方共同谈判，就有可能获得共同利益。

在多项议题的情况下，将双边谈判变成多边谈判会有点复杂。结果是，美国给予沙特阿拉伯协议数量的小麦，沙特阿拉伯给予日本协议数量的石油，而日本给予美国协议数量的汽车。如果后来美国违约，沙特阿拉伯却无法对美国报复；因为在这种情况下，沙特阿拉伯无法对美国停止提供任何物品，只可能中断对日本的石油输出。所以，实施多方协议可能会有这种问题。1946 年至 1994 年的关税及贸易总协定（General Agreement on Tariffs and Trade，GATT）及以后的世界贸易组织（World Trade Organization，WTO），都发现实施协议与处罚违反规则的国家是很困难的。

17.8 总　结

若能够达成协议，则可通过谈判分配剩余（即超额价值）。谈判可被当做各方一起找出解决方法并共同执行的合作博弈，或各方独自选择策略并达成均衡的非合作博弈。

纳什合作解以支付线性变化不影响结果、结果有效性、剔除无关因素对结果无影响等三条原则为基础。此解说明了这样一条规则，即各方获得的超过自身最低底线支付水平（也称 BATNA）的剩余的比例，是根据其相对的谈判实力而确定的。策略性地操纵最低底线可增加某一方的支付。

在轮流出价的非合作的规则之下，逆推法可以用来找出均衡点；一般而言，这会推导出这一结论即第一个回合的出价就被立刻接受。若剩余会随着出价被拒绝而递减，那么因其中一个参与人的拒绝而减少的总额就是另一个参与人均衡时获得的支付。若由于不耐烦而增加了延迟达成协议的成本，那么均衡时剩余的分配比例将与各方不耐烦的程度成反比。实验证明，博弈中参与人的出价往往高于达成协议所需的价格，这种结果被认为与参与人匿名参与博弈和相信公平有关。

谈判博弈中信息的非对称使得信号传递与信息甄别变得非常重要。有些参与人会传递信号来表明他们具有高 BATNA 或具有较强的耐性。另一些参与人则希望获得关于上述特性的真实信息。当愈来愈多的议题或谈判者被引入谈判时，达成协议便变得更为容易，但是谈判也会变得更有风险，实施也更困难。

关键术语

轮流出价（alternative offers）
谈判协议的最佳替代（best alternative to a negotiated agreement，BATNA）
递减（decay）
有效边界（efficient frontier）
有效结果（efficient outcome）
不耐烦（impatience）
纳什合作解（Nash cooperative solution）
剩余（surplus）
最后通牒博弈（ultimatum game）
可变威胁谈判（variable-threat bargaining）

习　题

1. 考虑康柏电脑公司与一个拥有网站 www. altavista. com 的加利福尼亚企业家间的谈判。最近接管了一家数码设备公司的康柏想要使用这位企业家的网站作为数码网路搜索引擎的地址：www. altavista. digaital. com。[15] 针对此网址，康柏和该企业家在 1998 年夏天举行了一场历时既漫长又艰难的谈判。虽然在此博弈中，这位企业家始终是处于弱势的参与人，但是最后却以 335 美元的价格达成协议。康柏在 8 月确认这笔交易，并在 9 月开始使用此地址，然而康柏却拒绝透露任何关于金钱的细节。在已知这些信息的情况下，请就下列项目提供意见：双方 BATNA 的可能值、它们的谈判实力或不耐烦程度，以及博弈是否达到了合作的结果。
2. A 与 B 正就 100 美元的分配进行谈判，A 先出价，提出这 100 美元在他们之间如何分配。如果 B 接受此出价，则博弈结束。如果 B 拒绝，那么总价值将减少 1 美元而成为 99 美元，然后轮到 B 出价。如此交替叫价，在一方拒绝对方出价后总价值就会损失 1 美元。A 的 BATNA 为 2.25 美元，B 的 BATNA 为 3.5 美元，那么博弈的均衡结果是什么？
3. 请考虑第 11 章习题 4 中比萨价格博弈的变形，其中一家店唐娜的厚薄饼的规模比另一家店皮尔斯的比萨大得多。请参考下面的支付表：

		皮尔斯的比萨	
		高价	低价
唐娜的厚薄饼	高价	156，60	132，70
	低价	150，36	130，50

非合作式的占优策略均衡为（高价，低价），所产生的利润唐娜的厚薄饼得到 132，而皮尔斯的比萨获得 70，总额为 202。若两者采取（高价，高价）策略，则其总利润将是 156＋60＝216，然而皮尔斯的比萨不同意这样的安排。假设若唐娜的厚薄饼另外付给皮尔斯的比萨一笔钱，两家店便可以达成都采取高价策略的协议。另一种解决办法就是非合作式的占优策略均衡。他们就协议内容进行谈判，唐娜的厚薄饼的谈判实力比皮尔斯的比萨大 2.5 倍。在最后的协议中，唐娜的厚薄饼要付给皮尔斯的比萨多少钱？

4. 假设有两个国家 E 国和 M 国正在就一项争议进行谈判。它们从 1 月开始每个月都要开会一次，轮流出价。假设总价值为 100。E 国政府正面临 11 月的改选，除非 E 国政府能够在 10 月的会议中达成协议，否则它将会败选。而败选对于 E 国政府而言，和在谈判中一无所获一样糟糕。而 M 国的政府并不在乎是否可以达成协议，也就是说，对于 M 国而言，无论是谈判延长或是开战，或是获得少于 100 的结果，都是一样的。这场谈判的结果将会如何呢？谁先行动会造成结果上的差异吗？

5. 根据第 4 题的答案，请讨论为什么实际谈判常会持续到最后期限。

6. 在一场不耐烦的轮流出价谈判中，当第一次出价时，假设 x 为参与人 A 所要求的总额，而 y 为参与人 B 所要求的总额。他们不耐烦的系数分别为 r 和 s。
   (a) 若使用近似的公式以 $x=s/(r+s)$ 表示 x，以 $y=r/(r+s)$ 表示 y，且 B 的不耐烦程度为 A 的两倍，则 A 会获得 2/3 的剩余，而 B 会获得 1/3 的剩余。请检验此结果的正确性。
   (b) 令 $r=0.01$，$s=0.02$，请比较下列两组的 x 和 y 值：使用近似公式 $x=s/(r+s)$ 和 $y=r/(r+s)$ 所求得的 x 和 y 值；使用由文中导出的公式 $x=(s+rs)/(r+s+rs)$ 和 $y=(r+rs)/(r+s+rs)$ 所得到的 x 和 y 值。

【注释】

[1] 在 5.3.2 节中，此博弈被用做支持纳什均衡不准确的论据。在涉及谈判的情形下，我们认为多重均衡点正式地表达出了之前所做的分析所声称的不确定性。

[2] John F. Nash, Jr.，"The Bargaining Problem," *Econometrica*, vol. 18, no. 2, 1950, pp. 155 - 162.

[3] 参见 Roger Fisher and William Ury, *Getting to Yes*, 2nd ed., New York: Houghton Mifflin, 1991。

[4] 只有一条凸向原点的曲线会和凹向原点的有效边界相切，在图 17—2 中，这条曲线为 c_2。所有比该条曲线低的其他曲线（如 c_1）都会和有效边界交于两点；所有比该条曲线高的其他曲线（如 c_3），都不会和有效边界相交。

[5] 若你已学习过基础微观经济学，你就了解经济学中生产可能性边界和社会无差异曲线的相切点所说明的社会福利最大化的概念。图 17—2 也具有类似的意义：谈判中的有效边界就如同生产可能性边界，而表征最大化目标的平行的曲线则类似于社会无差异曲线。

[6] 参见 Larry Debrock and Alvin Roth, "Strike Two: Labor-Management Negotiations in Major League Baseball," *Bell Journal of Economics*, vol. 12, no. 2, 1981, pp. 413 - 425。

[7] 为了让此论点简单一些，我们将这个过程想象为一对一谈判。实际上一定会存在着数位球迷与票贩，而整个状况会变成一个市场。我们在第 18 章中会对此市场进行分析。

[8] 为使得分析容易理解，我们忽略了球赛变得非常扣人心弦从而门票价值会随着节数减少而升高的可能性。不确定性让问题变得更复杂，但是获取处理这种可能性的能力的愿望应变成你学习更高级博弈理论的动机。

[9] 欲获得更多相关资料，请参考 Douglas D. Davis 与 Charles A. Holt，*Experimental Economics*，Princeton，N. J.：Princeton University Press，1993，pp. 263 - 269，以及 *The Handbook of Economics*，ed.，John H. Kagel and Alvin E. Roth，Princeton，N. J.：Princeton University Press，1995，pp. 255 - 274。

[10] 擅长博弈理论数学方面的学生（例如没有达成协议的组别），其表现都在平均以下，这可能是因为他们都努力获取额外的利润，所以都遭遇到了抵制，而女性的表现比男性略佳。

[11] 我们使用一种便捷的方法。我们简单地假设唯一的均衡支付是存在的。更严谨的理论会证明这些。想要获得更多有关这方面的资料，请参考 John Sutton，"Non-Cooperative Bargaining：An Introduce，" *Review of Economic Studies*，vol53，no. 5（October 1986），pp. 709 - 724。更严谨的理论（较难），请参考 Ariel Rubinstein，"Perfect Equilibrium in a Bargaining Model，" *Econometrica*，vol. 50，no. 1，1982，pp. 97 - 109。

[12] 有关此实务的观点，我们取自于 Andree Brooks，"Honing Haggling Skills，" *New York Times*，December 5，1993。

[13] 想要获得更充分的论述，请参考 Howard Raiffa，*The Art and Science of Negotiation*，Cambridge，Mass：Harvard University Press，1982。

[14] 经济学使用交换率或交换价格概念，此处用所交易的酒的瓶数与面包的条数来表示。重点是两国的可能利润会发生在任何介于 F 国刚好可以将面包交换成酒的交换率 2：1与 I 国可以进行同样行为的交换率 1：1 之间。在接近交换率2：1时，F 国会将其所有的 200 瓶酒拿来交换比 100 条多一点的面包，而不从事额外酒的生产；所以 I 国几乎获得所有的利润。相反，在接近交换率 1：1 时，F 国会获得几乎全部的利润。谈判的主题就是如何分配利润，因而双方就是要在该交换率或交换价格上进行谈判。

[15] 关于此博弈的详细资料，请参阅下列报道："A Web Site by Any Other Name Would Probably Be Cheaper，" *Boston Globe*，July 29，1998，以及 Hiawatha Bray，"Compaq Acknowledges Purchase of Web Site，" *Boston Globe*，August 12，1998。

第 18 章　市场与竞争

经济是关于生产和消费的体系。它的主要资源（如劳动力、土地和原材料等）用于生产商品和提供服务。这个体系的最后一个环节是商品和服务为最终消费者所购买和消费。这个体系随地点和时间的不同而不同，但它无论何时何地必须创造出将其组成部分连接起来的方式，即一套使不同供应者、生产者和买家相互进行交易的制度和安排。

在第 16 章和第 17 章，我们讨论了两种这样的制度：拍卖和谈判。在拍卖中，一般是一个卖者与几个现实的或者潜在的买者进行交易，虽然也存在相反的情况，即一个买者与几个现实的或者潜在的卖者进行交易，例如建筑工程或者供应项目的招标。谈判一般是一个买者面对一个卖者。

那么，余下的可能是所有经济制度中最常见的一种——市场（market），在市场里多个买方卖方同时进行交易。市场会让人联想起国外乡村集市的样子，商贩们有自己的货摊，买主来了就跟他们谈判。但大多数的市场却不这样。一个典型的城镇拥有两家或者更多的超市，镇上的居民在比较货源、价格、质量等因素后才决定在其中一家购买商品，或者在其中一家购买一些商品而在另一家购买另一些商品。一般来说，超市不会位于相邻的位置。尽管如此，它们都清楚必须相互竞争以吸引城镇居民的光顾，而这种竞争影响了它们对进什么货物、如何定价等问题的决策。商场和顾客之间不会有公开的谈判；价格都是标明了的，顾客要么支付这个价格，要么去其他地方购买。但是，另一家商场的存在限制了各家商场不能要价太高。

在本章中，我们将从策略角度研究市场。买卖双方在市场中应该如何行事？他们互动后达到什么样的均衡？在市场均衡时买者与卖者间的匹配是有效率的吗？

我们已经见过一些市场策略的例子。还记得第 4 章①中提到的两家餐馆的定价博弈吗？由于策略性互动引起两个卖者相互竞争，我们称之为双头垄断（duopoly，来自希腊语中的 duo 和 polein，分别表示“2”和“卖”的意思）。我们发现，这个博弈是个囚徒困境，两个卖者的定价都低于联合利润最大化时的价格水平。在第 11 章中，我们看到了一些方法，最主要的方法是重复进行博弈。借助于这些方法，两家餐馆可以解决囚徒困境并维持默契的高价共谋；于是就成为一个卡特尔（cartel）或者形成事实上的垄断（monopoly）。

但是，即使餐馆没有共谋，它们的价格也仍然高于其成本，因为每家餐馆都可能“俘虏”若干可供剥削并获取利润的客户。当市场中卖者数量多于两家，即卖者数量增加时，我们可以预料囚徒困境的情形会恶化，或者说卖者之间的竞争会加剧，其实就是卖者会定尽可能低的价格。如果有成百上千的卖者，其中的每个人就可能不会起到什么真正的战略意义上的作用；这时就该用更适用于大量独立小买者和小卖者的方法来取代博弈论方法。事实上，这一方法就是传统的供给与需求（supply and demand）的经济分析方法，它假定每个这样的小买者或小卖者都不能影响市场价格。每一方只决定在现行价格下买多少或者卖多少。然后，通过一些不可见的和非个人的市场对价格进行调整直到达到供求相等并使市场出清的水平。记得在第 2 章我们曾指出个人决策和策略互动博弈之间存在差异。传统的经济方法认为市场是个体决策者的集合，他们并不是通过策略而是以由市场决定的价格为媒介进行互动的。

现在我们更细致地讨论卖者和买者数量的增加如何加剧他们之间的竞争，从而由此产生一个不以个人意志为转移的市场，这个市场由市场价格调节供求达到平衡。换句话说，我们能将谈判和竞争中最基本的策略思想运用到我们的供求分析中。这种方法将加深你们对市场机制的理解，即使你们已在基础经济学课程和教科书中学过。如果还没学过，通过对本章的学习你们也会比那些学过经济学却没学过博弈论的人得到更透彻的了解。

本章除了对市场机制进行讨论外，还将指出竞争将会在何时以何种方式产生社会需要（socially desirable）的结果，并阐明何为“社会需要的”。最后，在 18.4 节和 18.5 节，我们考察了市场失灵时实现有效或公平分配的一些替代机制。

18.1 简单交易博弈

首先，我们讨论两人间的谈判，随后再研究更多人之间的谈判和竞争。设有两个人，一个卖者 S 和一个买者 B，就一套房子的价格进行谈判。卖者 S 可能是

① 原书如此，疑有误，应为第 5 章。——译者注

房子当前的所有者，从而获得相当于 100 000 美元货币价值的支付；也可能是一个建造者，其建造成本为 100 000 美元。无论哪种情况，该数量代表卖者的 BATNA，或者用拍卖的术语来讲，即保留价格。拥有该房子可使 B 可以获得 300 000 美元的支付，这是他愿意支付的最高价格。然而，现在 B 并未拥有一套房子，因此他的 BATNA 为 0。我们整节中都假设所有参与人都知道卖者的保留价格和买者的支付意愿；因此，谈判博弈不会遇到信息处理方面的附加的困难。关于信息处理方面的困境，我们将在以后的章节中连同市场机制一起进行讨论。

如果房子售价为 p，卖者获得 p 但是放弃房产；于是他的剩余（以千美元为单位）为（$p-100$）。买者拥有该房，这使他获得 300 的好处，而他为此支付了 p；于是他的剩余为（$300-p$）。两剩余相加得

$$(300-p)+(p-100)=300-100=200$$

这与特定值 p 无关。于是 200 为两方可通过进行一项交易实现的总剩余，即 S 和 B 相互达成交易时存在的额外价值。

交易中价格的功能是分配双方可获得的剩余。S 同意交易，他必定获得一个非负的剩余；因此 p 不小于 100。同样，如果 B 同意交易，则 p 不能大于 300。因此，谈判中 p 的取值范围应在这两个极限之间，但如果不对谈判的过程做更多具体说明，我们还是无法精确预测 p 的值。在第 17 章，我们谈到过两种对这个过程建模的方法。纳什合作模型在 S 和 B 的相对谈判能力给定的情况下，得到一种结果；轮流出价的非合作博弈模型在两者相对耐心程度给定的情况下得到另一种结果。

现在我们给出一种关于 BATNA、剩余和用第 17 章的方法得出的谈判范围的不同的几何解释。这种方法可帮助我们将建立在谈判基础上的那些思想与那些在市场中的交易思想联系起来。图 18—1 中，横轴表示数量，纵轴表示以千美元为单位计量的货币。曲线 S 表示卖者的行为，由三条或水平或竖直的黑线构成，

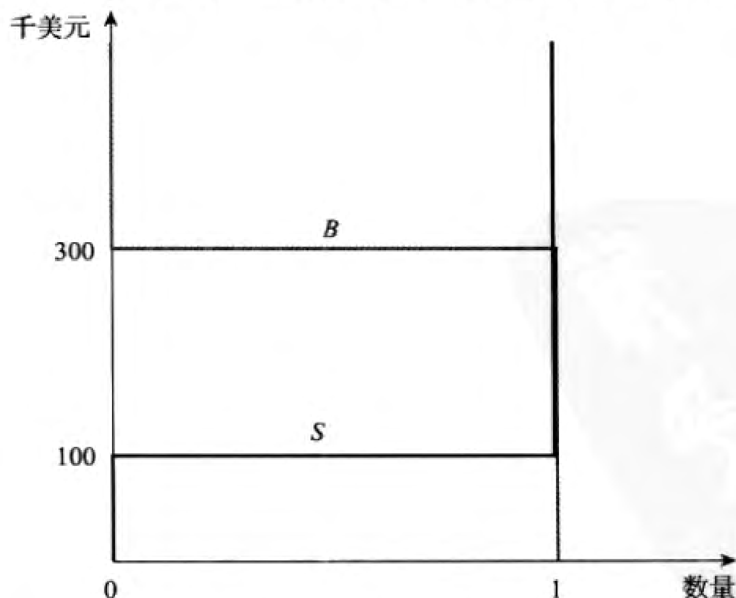


图 18—1　一个买者和一个卖者

形成一个向上的阶梯。[1]曲线 B 表示买者的行为，由三条或水平或竖直的灰线构成，形成向下的阶梯。

下面说明是如何画出卖者曲线的。首先，在纵轴上确定卖者保留价格的相应位置，这里是 100。然后在纵轴上从 0 到 100 画一条线，表明房子的价格不低于 100。接着，画一条长度为 1 的水平线（如图 18—1 黑线所示），这条线显示出卖者希望售出的房子的数量刚好是 1 套。最后，从这条线的右端画一条竖直向上的线，直到非常高的货币数量的区域。本例中的这条线强调没有其他卖者，因此无论对该房子出价多少，可供销售的数量也不会超过 1。

可以用相似的方法画出买者的谈判曲线 B。和前面一样，首先在纵轴上确定表示买者支付意愿的点，即 300。从这里画一条沿纵轴向上的线，表示对于高于 300 的价格，买者不愿意购买任何房子。然后，画一条长度为 1 的水平线（如图 18—1 灰线所示），表示买者希望购买的数量（1 套房子）。最后，从水平线的右端沿竖直向下方向画一条直到横轴的直线，意味着在这个例子中没有其他买者。因此不管索要的价格如何地低，对于超过 1 的数量不会找到任何顾客。

请注意两条阶梯形曲线 S 和 B 的竖直部分交叠，此时数量为 1，货币数量范围介于 100 和 300 之间。交叠部分预示买者和卖者之间的谈判范围。交叠线段的横坐标（这里为 1）就是交易的数量；交叠的纵向线段（这里处于 100 到 300 之间）代表最终谈判达成一致时的价格的范围。

你们中间学过经济学入门课程的人会发现曲线 S 就是卖者的**供给曲线**（supply curve），曲线 B 就是买者的**需求曲线**（demand curve）。对不熟悉这些概念的读者，我们可以在这里简单地解释一下。假设交易不是如图 18—1 中描述的二人博弈那样由买者和卖者自己发起并通过谈判达成的，而是发生在有组织的市场之中。一个中立的“市场交易商”（market maker）的角色就是集中买者和卖者并决定他们交易的价格，它能够决定可供售出的数量和每一个价格下的需求数量，然后调整价格并使市场出清，也就是说需求的数量等于供给的数量。市场交易商有时候是一个真实的人，比如金融市场中某一特定证券交易柜台经理，而其他时候只是经济理论中一个假设的概念。同样的结论对两种情况都适合。

接下来我们看看如何将市场的概念应用于我们关于谈判的例子的结果中。首先，考虑供给方。图 18—1 中的曲线 S 告诉我们，价格低于 100 时不会有房子出售，而价格高于 100 时只有 1 套房子出售。该曲线还表明，当价格恰好是 100 时，卖者既愿意卖出这套房子，也愿意作为最初的所有者而保留房子（或者如果他是建造者，他根本不修建这套房子）。于是可供销售的数量为 0 或者 1。那么，这条曲线表示出每一个可能的价格下可供销售的数量。在我们的例子中，介于 0 和 1 之间的小数数量没有任何实际意义，但是对于其他诸如小麦或者油料之类的被交易物品而言，我们很自然想到供给数量是连续变动的。

然后考虑需求方面。表示买者支付意愿的曲线即图 18—1 中的曲线 B，它表明对于高于 300 的任何价格，不会有任何买者愿意购买，但对于低于 300 的任何价格，有一个买者愿意购买刚好 1 套房子。在价格恰好为 300 时，买者购买与否同样开心；因此他会购买 0 或者 1 套房子。

需求曲线和供给曲线在数量为 1 且价格介于 100 到 300 的范围内重叠。当市场交易商在这个范围内选择任一价格时，就刚好出现一个买者和一个卖者。于是市场交易商让他们碰面并完成此项交易，市场就可出清。我们早先将这个结果解释为源于谈判达成的协议，但经济学家将其称为**市场均衡**（market equilibrium）。于是市场就是这样一种制度或者安排，它能带来与买者和卖者直接谈判相同的结果。此外，市场出清的价格范围也跟谈判中的范围是相同的。

市场制度使得每个买者和卖者被动地对由市场交易商制定的价格做出反应。当只有一个买者和一个卖者的时候这看起来很奇怪。事实上，市场的特征是竞争性的，这需要有多个买者和多个卖者。因此我们现在可以通过引入更多每种类型的行为人逐步扩展我们的分析范围，并在每一步都分析直接谈判与市场均衡之间的关系。

我们考虑这样一种情形，最初只有一个买者 B，他仍然愿意对一套房子支付最高 300 000 美元。但是现在我们引入第二个卖者 S_2，他拥有一套与第一个卖者（我们现在称其为 S_1）完全相同的房子。卖者 S_2 的 BATNA 为 150 000 美元。这可能比 S_1 的要高一些，因为作为房子的拥有者 S_2 认为在他房内居住的主观价值比 S_2 更高，或者作为建造者 S_2 的建造成本比 S_1 更高。

S_2 的存在意味着 S_1 不能指望可以以高于 150 的任何价格与 B 成交。比如，如果 S_1 希望与 B 以 152 的价格交易，S_2 就会降价，向买者提出 151 的价格。S_2 仍然会得到正的剩余，而不会被淘汰出局而一无所获。同样，S_2 不可能指望要价会超过 150，因为 S_1 甚至可能更热衷于降价。于是竞争性卖者的存在将这个博弈中谈判的范围由（100，300）缩小为（100，150）。均衡交易仍发生在最初两个参与人 S_1 和 B 之间，因为它将发生在介于 100 和 150 之间的某个价格上，而这个价格范围会使 S_2 退出交易。于是第二个卖者的存在压低了最初那个卖者所能得到的价格。价格定格于 100 到 150 之间哪个位置取决于 B 和 S_1 之间的相对谈判能力。即使 S_2 在低于 150 的任何价格下退出，如果 B 由于其他这里未提及的理由（比如相对耐心程度）而拥有巨大的谈判能力，于是他可能会拒绝 S_1 提出的接近 150 的价格，并坚持一个更接近 100 的价格。如果 S_1 的相对谈判能力较弱，他可能就不得不接受这样一个价格。

我们将这一情形中的需求曲线和供给曲线表示在图 18—2（a）中。需求曲线和之前相同，但是供给曲线有两个台阶。对低于 100 的任何价格，任何卖者都不愿出售，因此数量为 0；在纵轴上是 0 到 100 之间的一条竖直线。对 100 到 150 之间的任何价格，S_1 愿意出售而 S_2 不愿意。于是提供给市场的数量为 1，在这个数量上，竖直线由 100 延伸至 150。对高于 150 的任何价格，两位卖者都愿意出售，数量为 2；在这个数量上，竖直线从价格 150 开始向上，这是供给曲线的最后一段。我们再在价格为 100 时在数量 0 到 1 之间以及价格为 150 时在数量 1 到 2 之间分别画线段，使供给曲线为一个不间断的整体。

需求曲线和供给曲线的交叠部分现在是一条代表数量为 1，价格范围为 100 到 150 的竖直线段。这恰好是我们先前考察谈判而非市场时发现的谈判范围。于是，两种方法预示了同样的结果范围：转手 1 套房子，保留价格低的卖者能售出，最终价格介于 100 000 美元和 150 000 美元之间。

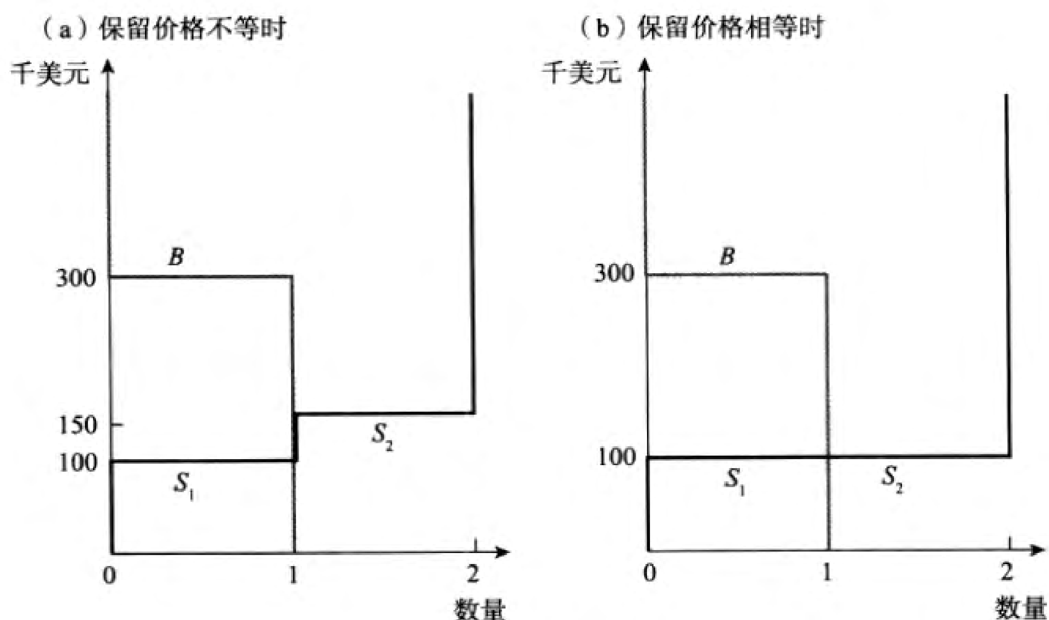


图 18—2　一个买者，两个卖者

如果两位卖者保留价格相同，都为 100 000 美元呢？在这种情况下，谁也不会以高于 100 的价格出售。比如，如果 S_1 要价 101 000 美元，那么 S_2 面对根本无法售出的前景，就会只要 100 500 美元。谈判的范围就会完全消失。其中一个卖者会成功地以 100 000 美元卖出他的房子。另外一位会退出继续持有，但在这个价格上，他卖出房产与否都同样开心。

我们将在市场供给和市场需求背景下的上述情形用图 18—2（b）表示出来。任何一位卖者在低于 100 的价格下都不出售他的房产，而只要价格高于 100，两位卖者都会出售各自的房产。于是供给曲线在高度为 100 时的水平线段长度为 2，最后右端是无限长的向上的竖直线。需求曲线不变，于是两条曲线仅在点（1，100）相交。现在价格范围缩减为一个数值，这个价格代表市场均衡：转手了 1 套房子，其中一位卖者（哪一位并不重要）售出房子，价格恰好为 100。两位保留价格相同的卖者之间的竞争阻碍了其中任何一位的收入超过保留价格；买者以高额剩余的形式享受了卖者之间的竞争带来的利益，即 300－100＝200。

反之亦然，即只有一个卖者，但有第二个买者 B_2，他愿意对一栋房产支付 200 000 美元。在这样的情况下，第一个买者（记为 B_1）不可能指望在低于 200 的价格下得到房子；因此，如果 B_1 要获得房子，他将不得不支付高于这一数量的价格。S 和 B_1 达成一致的可能价格范围现在为（200，300）。这一范围又比最初只有一个买者和一个卖者时的（100，300）更加狭小。至于用谈判方法和市场供求曲线进行精确分析，我们留给读者来做。

我们继续讨论谈判和市场，在这里我们要引入更多的买者和卖者。下面我们讨论每种类型的交易者各有两人的情形。假设卖者 S_1 和 S_2 的保留价格分别为 100 和 150，买者 B_1 和 B_2 的最高支付愿意分别为 300 和 200。我们可以预计，博弈中

的谈判范围为（150，200），可证明如下。

每位买者和每位卖者都为自己寻找最优交易。他们可以提出或者接受一项试探性的提议，与此同时继续寻找更好的，直到无人找到更好的时为止，在这个点上整个交易终止。考虑这样一个试探性的提议，其中一套房子（比方说是 S_1 的）在价格 $p_1 < 150$ 时出售，而另一套（S_2 的）在价格 $p_2 > 150$ 时出售。于是愿意支付更高价格的买者，比方说 B_1，就会找到 S_1，并向他提出（$p_1 + p_2$）/2 的价格。这一价格高于 p_1，因此对于 S_1 而言好于最初的试探性的提议，同时低于 p_2，因此对 B_1 而言也是如此。这个价格对其他两个交易者 B_2 和 S_2 而言则较糟糕，但是由于 B_1 和 S_1 可自由选择他们的交易对象进行交易，因此他们无计可施。于是任何低于 150 的价格不可能成交。类似地，你可以验证一下任何高于 200 的价格同样不可能成交。

所有这些只是对我们先前推论过的只有少量交易者的情形的细微的扩展。但是在每种类型交易者都有两人的情形中，出现了一个有趣的新特征：两套房子（假定它们同质）都只能在相同的价格出售。若非如此，计划在最初提议中支付更高价格的买者就会与要价更低的卖者达成一项对双方都更有利的交易。

注意这里竞争与谈判之间的本质上的区别。你是与“交易的另一方”的某个人进行**谈判**；而你是与“交易的同一方”的某个人进行**竞争**，以获得机会与交易的另一方的某个人达成交易。买者与卖者**谈判**，买者与其他买者**竞争**，卖者与其他卖者**竞争**。

我们在图 18—3（a）中画出有两个卖者和两个买者参与的供给曲线和需求曲线。现在，每条曲线由三条竖直线段和两条水平线段构成，水平线段均在供求双方保留价格的那一点上与竖直线段相连。曲线在数量为 2 且价格范围为（150，200）的部分交叠。交叠部分就是市场均衡所处的位置。两套房屋都会以同样的价格转手；而到底是 S_1 卖给 B_1，S_2 卖给 B_2，或是 S_1 卖给 B_2，S_2 卖给 B_1，无关紧要。

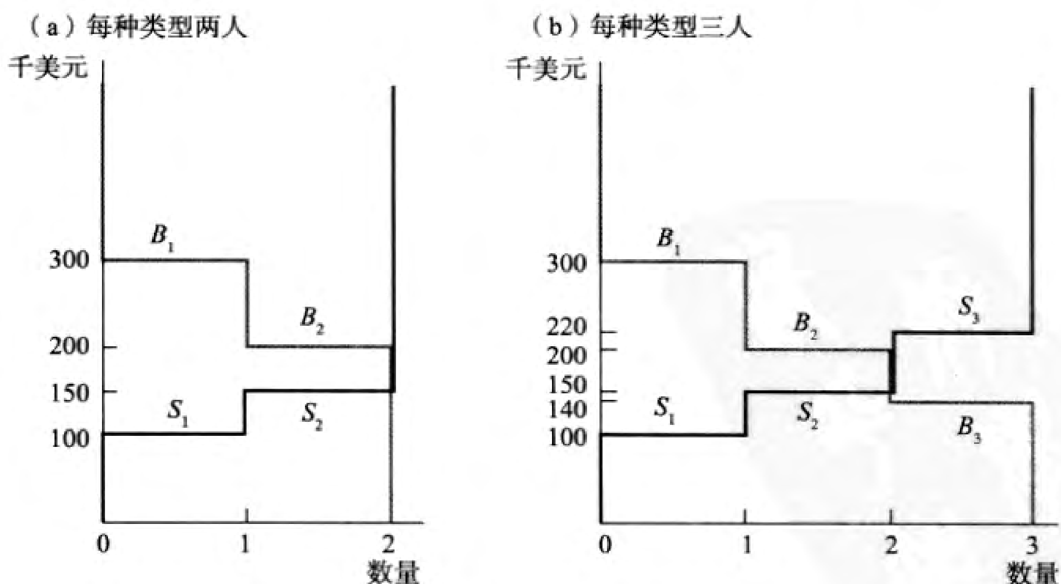


图 18—3　更多买者和卖者

最后，我们探讨三个买者和三个卖者的情形。假设最初的两个卖者 S_1 和 S_2 的保留价格不变，分别为 100 和 150，而第三个卖者 S_3 的保留价格更高，为 220。同样，最初两个买者的支付意愿如前不变，分别为 300 和 200，而第三个买者的支付意愿较低，仅为 140。新的买者和卖者根本不会影响谈判或者其范围，即仅有 2 套房子在卖者 S_1、S_2 和买者 B_1、B_2 之间达成交易，交易价格仍介于 150 与 200 之间。B_3 和 S_3 未能达成交易。

来看看 B_3 和 S_3 为何在交易中出局。首先假设 S_3 即将达成某试探性协议要出售他的房子，但仅有 B_1 可能成为买者，因为无论 B_2 或者 B_3 都不愿支付 S_3 的保留价格。假设协议的价格是 221，则 B_1 得到的剩余为 79，而 S_3 仅仅得到 1。然后考虑其他人必须如何配对。S_1 的保留价格为 100，必须卖给支付意愿为 140 的 B_3；S_2 的保留价格为 150，必须卖给支付意愿为 200 的 B_2。S_1 卖给 B_3 的剩余为 40，S_2 卖给 B_2 的为 50。因此 S_1 可从销售中最多得到接近 40 的剩余，S_2 最多可得到接近 50 的剩余；任何卖者都可通过找到 B_1 并达成一项对双方都有利的替代交易而获利。比如，S_2 可找到 B_1 说，“你愿意支付的价格为 300 而我的保留价格为 150，我们之间有总数为 150 的可利用剩余。你在与 S_3 的交易中得到 79；我在与 B_2 的交易中得到的不足 50。我们进行交易，你可以得到 90 而我获得 60，这对我们大家都更有利。”B_1 会接受这个交易，而与 S_3 的试探性协议也会化为泡影。

我们例子中推导所用的数量都是特定的，但论证过程完全具有一般性。对任意数量的买者和卖者，首先按保留价格对卖者递增排序，按意愿支付价格对买者递减排序。然后将相同序号的买者和卖者配对。当你发现第 n 个买者的意愿支付的价格低于第 n 个卖者的保留价格时，你就知道了所有序号为 n 的买者和卖者不会达成交易。

如图 18—3（b）所示，使用市场供求曲线分析法，论证会更加清晰。供求曲线的画法如前，我们这里就不再具体阐述了。两曲线交叠于数量 2、价格范围为从 150 到 200 的地方。因此，市场均衡时，有 2 套房子以两个极限范围内的某个价格转手。给定可能的价格范围，我们可知道 B_3 的支付意愿 140 将低于市场价格，S_3 的保留价格 220 将高于市场价格。因此，正如我们在一般情形中所言，第 n 个买者和卖者不能达成交易，我们可以断言这两个参与人在市场均衡中不起作用。同样地，S_1 卖给 B_1、S_2 卖给 B_2，或者 S_1 卖给 B_2、S_2 卖给 B_1，都不重要。

为了让这一描述更清楚明确，假设上述情形中的市场价格为 175。那么交易者的剩余：S_1 为 75，S_2 为 25，B_1 为 125，B_2 为 25，S_3 和 B_3 为 0；总的剩余为 250。你可以检验一下，当范围介于 150 至 200 时，无论价格多少，总的剩余都相同。

难道我们就没有更好的配对方式取得更好的结果吗（尤其是三套房都被售出）？答案是否定的。谈判过程或者市场机制产生了最大可能的剩余。和前面一样，为了提出一个替代解决方案，我们强行用 B_1 和 S_3、B_3 和 S_1、B_2 和 S_2 配对。那么 B_1 和 S_3 配对产生的剩余为 300－200＝80①，B_3 和 S_1 配对产生的剩余为 140－100＝40，最后剩下的 B_2 和 S_2 配对产生的剩余为 200－150＝50。这个替代

① 原文如此，疑有误。——译者注

方案中总的剩余加起来仅有170，比谈判或者市场均衡时产生的总剩余250少了80。问题在于B_3只愿意支付140，一个低于均衡范围最低价格的价格，但是S_3的保留价格为220，高于均衡范围的最高价格；在他们之间，博弈后得到一个负的剩余为140－220＝－80。如果他们有机会进行交易，就无法回避总的剩余将减少80这样一个事实。

虽然通过谈判或者市场均衡产生了最大的总剩余，但它还是不能对剩余的平均分配提供必要的保证，即B_1、B_2、S_1和S_2瓜分总的250而B_3和S_3一无所获。在我们的替代方案中，每个人都得到一些，但是配对过程不允许重新谈判。如果你对公平足够关注，你可能更偏好替代方案，而不管所产生剩余的水平。你可以对重新谈判的过程进行限制以产生你偏好的结果。最好的是，你可以进行谈判或让市场机制发挥作用以产生最大剩余，然后从享受剩余的人中拿走一些剩余，利用税收—转移支付政策对其他人进行二次分配。公共经济学对这些问题有深入探讨。

我们这里姑且暂停，概括一下本节的内容。我们已经找到了一种思考买者和卖者间谈判如何推进以确保双方达成共赢交易的方法。我们还将谈判过程的结果与基于供求分析的市场均衡的结果联系了起来。虽然我们分析的只是仅涉及少量买者和买者且数据简单的例子，但是其思想具有普遍性。我们将在后面几节以更具一般性的方法对此加以阐述。在这个过程中，谈判的概念已经超出了交易的范畴。正如我们在17.7节看到的，当存在多个参与人、有多种商品和服务可供交易时，最优交易可能是多边的（multilateral）。一种真正的一般性理论不仅适用于二人交易，而且也适用于三人交易或多人交易。这样的团体，有时称为联盟（coalition），能在单个人和单个组合寻找更好的替代办法时，集合起来进行临时交易，直到单个人或单个组合不能通过谈判获益时为止。于是临时交易就成为最终交易并成交。下一节我们继续讨论关于达成互利交易的问题。

18.2 核

为分析以达成互利交易为目的的谈判和再谈判过程，我们现在引入一些符号和一种一般性理论，随后我们将其应用于一个简单的例子。我们将发现，由于参与人数量的增加，联盟形成、再形成和交易达成的复杂过程所产生的结果与传统市场所产生的结果非常接近。于是我们就可以看到市场就是一种可以简化数量巨大的交易者之间达成交易的复杂过程的制度。

考虑n个人参与的一般性博弈。n个参与人简单记为1，2，…，n，所有参与人的集合为$N=\{1, 2, 3, \cdots, n\}$。集合N中的任何一个子集为一个**联盟**C（coalition）。当有必要详细指明一个联盟的所有元素时，我们用大括号列出；$C=\{1, 2, 7\}$就可能是这样一个联盟。我们允许N是自身的一个子集，并且称之为**大联盟**G（grand coalition）；于是$G=N$。用组合数学的一个简单公式可得到N的只包括n个元素中的1个，或者2个，或者3个……或者n个时的全

部子集的个数。可能的子集个数为（2^n-1），于是总共有（2^n-1）个可能的联盟。

18.1 节讨论的简单二人交易博弈中，$n=2$，于是我们可以想到存在 3 个联盟：2 个单成员（分别是 1 个买者和 1 个卖者）联盟，和 1 个由买者和 1 个卖者共同构成的大联盟。当有 2 个卖者和 1 个买者的时候，我们有 7 个即（2^3-1）个联盟：3 个单成员联盟；3 个成对联盟，即 $\{S_1, B\}$，$\{S_2, B\}$ 和 $\{S_1, S_2\}$；1 个大联盟 $\{S_1, S_2, B\}$。

一个联盟能在其自身成员之间达成交易以获取所有可能的共同利益。对任一联盟 C，我们令函数 $v(C)$ 为其成员可以获取的总剩余，而不考虑该联盟以外的参与人（即集合 $N-C$）的行动。这个剩余的数量被称为联盟的**安全水平**（security level）。当我们说"不考虑联盟 C 以外的参与人的行动"时，我们设想一下最坏的情形，即非联盟 C 的参与人采取了对于联盟 C 而言最坏的行动时，联盟 C 中的成员会发生什么情况。这种思想与我们在第 7 章和第 8 章已经学习过的个体参与人的最小最大（值）思想是相似的。

对 $v(C)$ 的计算来源于博弈的具体细节。我们以第 1 节提到的房产交易的简单博弈为例。离开另一方成员的合作，任何参与人自身都不能获取剩余。于是所有单成员联盟的安全水平为 0，即 $v(\{S_1\})=v(\{S_2\})=v(\{B\})=0$。联盟 $\{S_1, B\}$ 可实现 300－100＝200 的总的剩余，S_2 不能采取任何行为减少二人的剩余（纵火是博弈规则所不允许的）；于是 $v(\{S_1, B\})=200$。类似地，$v(\{S_2, B\})=300-150=150$。两个卖者在没有买者的情况下，自身不能获得任何额外支付，因此 $v(\{S_1, S_2\})=0$。

对一个博弈中任何一个联盟 C，我们都可找到它的安全水平 $v(C)$。$v(C)$ 也被称为该博弈的**特征函数**（characteristic function）。我们假定特征函数是共同知识（common knowledge），不存在非对称信息。

接下来，我们探讨博弈可能的结果。**配置**（allocation）就是对参与人剩余数量的一个排列（x_1，x_2，…，x_n）。如果配置在逻辑上与博弈规则相符，即它来源于按规则成交的交易，那么配置就是**可行的**（feasible）。任何可行的配置会被作为一项试探性的交易提出，但当某个联盟解散对自身更有利时，该交易就会被推翻。我们说一项可行的配置被一个联盟 C **阻挡**（block），如果

$$v(C) > \sum_{i\text{在}C\text{中}} x_i$$

这个不等式表示联盟的安全水平（即它能向其成员确保的总剩余的最小值）超过其成员在提议的试探性配置中获得的剩余的总和。如果该不等式成立，那么这个联盟就能形成，并将在成员中展开谈判使得他们中所有人比在提议的配置中的情况更好。因此，联盟暗藏的解散的威胁（如果获得的剩余并不足够高的话，就反而会结束交易）是可信的。这就是联盟如何能推翻（阻挡）提议的配置的。

不能被任何联盟阻挡的配置的集合包含所有可能的稳定的交易，或博弈的谈判范围。这个集合不会由于任何一个组合寻求更好的交易而被进一步减少。我们称这个配置的集合为博弈的**核**（core）。

18.2.1 数字例子

在 18.1 节中我们已经得到简单二人交易博弈中的阻挡，虽然我们并未使用这一术语。然而，我们现在可以看到我们称为“谈判范围”的东西正好对应于核，因此我们可以将结论重新表述如下：在仅有一个买者和一个卖者的情形下，谈判范围是从 100 到 300。这个结果对应于一个核，它是由介于买者获得 200 剩余，卖者获得 0 剩余（价格为 100）和买者获得 0 剩余，卖者获得 200 剩余（价格为 300）之间的所有配置构成的。

现在我们使用新的分析工具来考虑一个支付意愿为 300 的买者遇见两个保留价格均为 100 的卖者的情形。三个参与人之间博弈的核是什么呢？

考虑一种剩余的配置，x_1，x_2为两个卖者的剩余，y 为买者的剩余。对核中的这种配置，没有任何一个联盟其成员所得的会比此处更多。我们知道每个单成员联盟只能获得 0，因此一个**核配置**（core allocation）必须满足条件

$$x_1 \geqslant 0, \quad x_2 \geqslant 0, \quad 和\ y \geqslant 0$$

接下来，一个买者和一个卖者的联盟能获得总共为 300－100＝200 的剩余，但两个卖者的联盟只能获得 0。因此，若使提议的配置在被一个二人联盟阻挡时能有效，必须满足

$$x_1 + y \geqslant 200, \quad x_2 + y \geqslant 200, \quad 和\ x_1 + x_2 \geqslant 0$$

最后，全部三个参与人总共可以获得的剩余为 200，因此，如果提议的配置未被大联盟阻挡，我们必定有

$$x_1 + x_2 + y \geqslant 200$$

但由于可获得的总剩余仅为 200，$x_1 + x_2 + y$ 不可能大于 200。因此我们得到方程 $x_1 + x_2 + y = 200$。

给定最后这个方程，我们现在就可以证明 x_1必定为 0。如果 x_1严格为正，并且将 $x_1 > 0$ 与二人不等式之一的 $x_2 + y \geqslant 200$ 结合起来，我们就可以得到 $x_1 + x_2 + y > 200$，而这是不可能的。类似地，我们还发现 $x_2 = 0$。接下来就得到 $y = 200$。换言之，卖者得不到剩余，而全部剩余给了买者。为了达到这个结果，其中一套房子必须按其保留价格 $p = 100$ 出售。至于是哪一套，无关紧要。

下面我们通过对竞争的直觉（intuition）展开分析。当两个卖者的保留价格相同时，他们之间的竞争就会变得非常激烈。如果其中一个在价格为 102 时对一项交易进行谈判，另一个就会准备降价而向买者报价 101，这比没有交易强。这个过程中，卖者“互相压价”（beat each other down）直到 100。本节中，我们已经研究了关于核的一般性理论，并用数学方法对之进行了解读，从而得到了一个与直觉相符的结果。我们希望数学推导与直觉思考这两种方法都会给你留下深刻的印象。当问题变得更加复杂时，更多的对象、更多的买者、更多的卖者会使纯粹的直觉思考更加困难。数学为我们提供了解决所有这一类问题的一种一般性算

法，你将有足够多的机会使用它。但是当用数学解决了更复杂的问题后，你也该停一下，将结论与直觉联系起来，就像我们这里所做的一样。

18.2.2 核的一些特征

1. 有效性

前面我们看到，如果不被**任何**联盟（包括大联盟）阻挡，配置就在核中。因此，如果一个核配置中 n 个成员的剩余为（x_1，x_2，…，x_n），这些剩余之和至少必须与大联盟可获得的安全水平一样多，即

$$x_1+x_2+\cdots+x_n\geqslant v(\{1,2,\cdots,n\})$$

但大联盟可获得任何可行的剩余配置。因此，对**任何**可行的剩余配置，比如（z_1，z_2，…，z_n），必有

$$v(\{1,2,\cdots,n\})\geqslant z_1+z_2+\cdots+z_n$$

这两个不等式一起暗示出，要找到另外一种使每个人得到比核配置更高的支付的可行的配置是不可能的。我们**不可能**同时有

$$z_1>x_1,\quad z_2>x_2,\cdots,\quad 和\ z_n>x_n$$

由于这个原因，核配置就是一个**有效配置**（efficient allocation）。用经济学家的行话来说，就是**帕累托有效**（Pareto efficient），这一术语是以最早引入有效性概念的维尔弗雷多·帕累托（Wilfredo Pareto）的名字命名的。

有效性听起来不错，但它只是我们希望核配置具有的好的特征之一。有效性只是说其他的配置都不能同时改善每个人的支付；核配置仍可能会使得一些人的支付非常低而一些人的支付非常高。这就是说，有效性并不代表核配置具有均等性、公正性和公平性。在我们的交易博弈中，如果存在另一个具有相同保留价格的卖者，则此卖者获得的剩余将为 0；如果没人与其保留价格相同，他就有可能获得正的剩余。这大概是运气，而非公平。我们还看到一些保留价格很高的卖者和一些支付意愿很低的买者之所以被逐出交易，是因为其他人可获得更多的总剩余；这同样是为获得有效性而牺牲了均等性。

核配置总可产生最大总剩余，但它并不能调节剩余的配置。在两个买者、两个卖者的例子中，较高支付意愿的买者（B_1）获得更多剩余，这是由于他与另一个买者（B_2）的竞争意味着两人支付同样的价格。类似地，低保留价格的卖者（S_1）获得大多数剩余，或者等价地，低成本的建造者（S_1）得到大部分的利润。

在我们的简单交易博弈中，核一般是一个范围而非单纯一个点。类似地，其他博弈中的核也可能有多种配置。所有的配置都是有效的，但它们使得参与人之间的总支付的分配各不相同，一些比其他的更公平。有效性的概念未涉及存在多个核配置时，我们应如何选择某个特定的核配置。这需要另外的理论或者分析方法。在本章后面的部分，我们将简单考察一下关注参与人之间支付分布的其他机制。

2. 存在性

能保证所有的博弈都有一个核吗？不能。当博弈中某一些参与人对别的参与人有负的溢出效应时，一个由互相伤害的参与人组成的联盟就很有可能推翻每一个临时提议的配置。这样的核可能是空的。

一个极端的例子就是**垃圾博弈**（garbage game）。有 n 个参与人，每人有一袋垃圾。每人既可将垃圾倒在自己的院子里也可将垃圾倒在邻居的院子里。他甚至可以将垃圾分开，一些倒在自己的院子里，一些倒在邻居的院子里。有 b 袋垃圾倒在自己院子里的支付为 $-b$，其中 b 不必为整数。于是 $m<n$ 时，一个 m 人的联盟的安全水平为 $-(n-m)$，因为他们将自己的 m 袋垃圾倒在非联盟成员的院子里，但是在安全水平计算显示的最糟糕的情形下，所有（$n-m$）个非联盟成员会将他们的垃圾倒在联盟成员的院子里。于是博弈的特征函数为 $v(m)=-(n-m)=m-n$，其中 $m<n$。但是大联盟中没有非联盟成员，也无法将垃圾倒在他们的院子里，于是 $v(n)=-n$。

现在，很容易看到任何提议的配置都可能被阻挡。最严格的验证就是 $(n-1)$ 个人的联盟。它具有高达 -1 的安全水平，因为他们将把所有的垃圾都倒在被联盟所排斥的那个不幸的人的院子里，而全部（$n-1$）个人只需承受他 1 袋垃圾。假设第 n 个人当前被排斥，于是他可以找到联盟的任意（$n-2$）个成员，比如后面的（$n-2$）个，向他们提出一项甚至更好的交易。新的（$n-1$）人的联盟（即从 2 到 n）将把垃圾倒在第 1 个人的院子里。第 n 个人将接受第 1 个人的整袋垃圾。于是新的（$n-1$）人联盟中其余人的境遇甚至比在最初的（$n-1$）人联盟中时还要好，因为他们之间分享的剩余为 0 而不是 -1。第 n 个人由于草坪上只有 1 袋垃圾，比起最初他作为被排斥的人而要承受 $(n-1)$ 袋垃圾的情形来，境遇显著改善了。于是最初的配置就阻挡了。对于每一个配置而言都是如此，因此核不存在，或说是空的。

这个问题产生于产权（property rights）的性质，这一点并未在博弈中得到明确说明。人们有权在别人的院子倾倒垃圾，就像有时候在有些地方，企业有权污染大气和水，人们有权吸烟。在存在“污染者权利”的制度下，负的外部性不受限制，联盟形成与变化的过程不稳定，从而产生空核。在另一种制度下，如果人们有权避免其他人产生的污染，这个问题就不存在。在我们的博弈中，如果产权改变以至于未经他人许可无人能向其他人倾倒垃圾，那么核就是非空的，每个人将垃圾留在自己院子里并得到 -1 的支付的配置不可能被阻挡。

18.2.3 讨　论

1. 合作博弈或非合作博弈？

核的概念是用来对博弈进行抽象描述的，它是从特征函数开始的。关于一个联盟 C 如何获取等于其安全水平 $v(C)$ 的总剩余的问题，这一概念并未提供具体细节。这些细节对于每个实例都是特殊的。同样，当我们找到一个核配置时，产生该配置的单独行为是如何实施的也并未得到说明。这里隐含地假设了存在某个

外部裁判确保所有参与人都扮演被分配的角色并采取被分配的行动。因此核必须理解为合作博弈理论范围内的一个概念，行为是联合选择和实施的。

但是独立参与人可自由选择加入联盟或者解散现有联盟并成立新的联盟，这为想象中增加了很大的非合作成分。这就是为什么核常常是思考竞争的一种好方法的原因。然而，要使竞争能自由地展开，所有买者和卖者享有平等和充分的自由以形成和重新形成联盟是很重要的。如果只有卖者是博弈中起作用的参与人，买者只能被动接受他们制定的价格且不能寻求更好的交易，那么情况会怎么样？于是最大化总利润的卖方博弈的核可以形成卖者的大联盟，即卡特尔或者垄断。现实中经常可以看到这类现象。在任何时间，市场上都有大量的买者，但由于每个单独的买者非常小并且不常进入市场，他们聚在一起或者与卖者谈判以求更好的交易是非常困难的。卖者的数量少，而且卖者之间有重复的关系，他们更有可能形成联盟。这就是为何反托拉斯法试图防止卡特尔的形成并保持竞争，然而几乎不存在防止买方联合的类似法律。[2]

2. 附加的和非附加的支付

我们将联盟的安全水平定义为向其成员所确保的全部支付之和的最小值。当我们用这种办法增加单个成员的支付时，我们将做两个假设：(1) 支付使用一些可比较的单位来度量，货币是最常用的标准；(2) 联盟以某种方式解决了内部的谈判问题。甚至当总支付很高的时候，如果一些成员得的多而其他人得的少，联盟仍然可能解散，除非前者对后者做出补偿，使得每人在联盟中比在联盟外境遇更好。这样的补偿必须采取支付可以度量的转换单位的形式，货币再一次具有代表性。

如果货币（或者其他一些所有人都有需求的并用对于所有人来说都是可比较的单位度量的事物）不是可转移的，或者如果人们的支付是货币数量的非线性函数（比如由于风险规避），那么联盟的安全水平就不能单独用其总支付描述。于是我们不得不明确全部成员各自的支付。换言之，特征函数 $v(C)$ 不再是一个数字，而是一组列举出所有成员支付组合的向量，这些支付组合对本联盟是可能的，可不必考虑其他的联盟。这种情形下——技术上称为**非转移效用**（nontransferable utility）的情形——的核的理论，是可以发展出来的，并且你将在对合作博弈理论[3]的更高深的研究中更多地看到这种理论。

18.3 市场机制

我们已经考察过交易者数量不大时的市场供给曲线和需求曲线。将其一般化以符合更为现实的市场状况是很容易实现的，如图 18—4。每种商品都可能有数百或者数千个潜在的买者和卖者；就市场的每一方而言，他们都可以测出保留价格和支付意愿的范围。对于纵轴测度的每个价格 p，标有 S 的曲线表示保留价格低于 p（即他们愿意以此价格出售）的所有卖者提供出售的商品的数量。当我们将 p 增加到足够大，超过了一个新的卖者的保留价格时，这个卖者的商品的数量就进入

市场，于是供给数量增加。我们将这个数量表示为 p 取相应值时曲线的水平部分。经这个过程产生的向上的楼梯形状图形是供给曲线。类似地，买方的向下的楼梯状图形表示每个价格下买者愿意购买（需求）的数量，因此称为需求曲线。

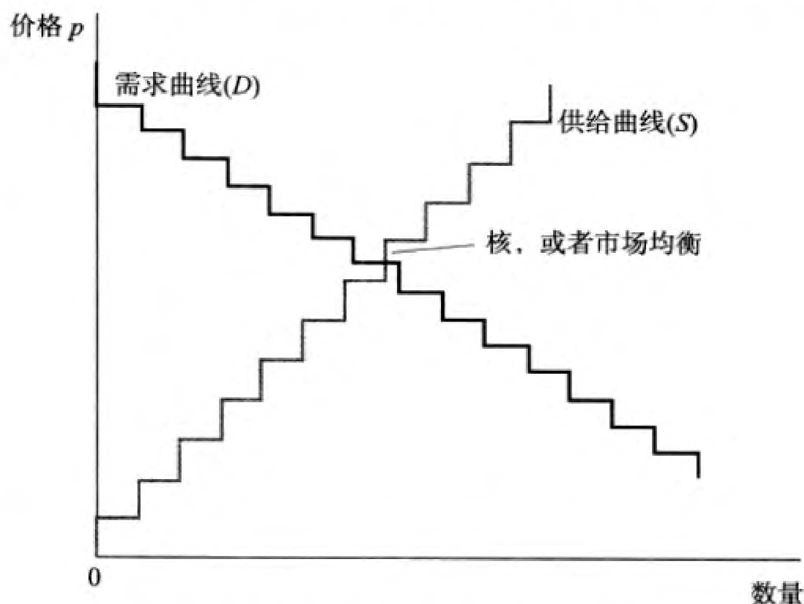


图 18—4　多个买者和卖者时的市场均衡

如果所有的卖者都有相同的产出投入和相同的生产技术，他们就可能有相同的成本从而相同的保留价格。这种情况下，供给曲线可能是水平的而非阶梯形的。然而，一般来说，买者之间存在着特殊的偏好差异，因此需求曲线总是阶梯形曲线。

市场均衡产生于两条曲线的交点，相交有三种可能的方式。第一种是在某一个数量竖直交叠。该交叠部分意味着这个数量恰好是交易量，交易价格（对所有交易者都一样）则可为交叠部分范围内的任意值。第二种可能性就是两条曲线某段水平部分交叠。此时价格是唯一决定的，但数量可在交叠所示范围内变化。水平交叠的特性意味着单一的价格同时是一个或多个卖者的保留价格和一个或多个买者的支付意愿；在这个价格下，进行交易与否对两种类型的交易者无差别，因此交易数量位于交叠部分具体哪个位置并不重要。第三种可能性是一条曲线的水平部分与另一条曲线的竖直部分相交。此时两条曲线只有一个交点，价格和数量都是唯一决定的。

每种情况下，当中立的市场交易商参与进来时，这个价格（即使不是唯一决定的）就是需求数量与供给数量相等、市场出清时的价格。任何这样一个使市场出清的价格都被称为市场均衡价格。市场交易商的目的就是要找到这样一个价格，他是通过试探和纠错的过程寻找这一价格的。他喊出一个价格，提出买和卖。所有这些提议最初都是试探性的。如果当前价格下需求的数量高于供给的数量，他就会提出一个稍高的价格。市场交易商通过这种办法调整试探性价格直到市场出清。只有当买和卖的临时提议切合实际时，它们才会在市场均衡价格下实现。

当成百上千的买者依据其支付意愿排序，成百上千的卖者依据其保留价格排序时，就买卖双方任一方而言，他们相继达成的交易之间的差别都非常小。因

此，需求曲线的每一个阶梯只下降一点点，供给曲线也只上升一点点。这样的需求曲线和供给曲线，即使有一段垂直的重叠部分，在市场均衡处，价格不确定的范围也是非常小的，市场出清价格几乎是唯一的。

我们通过简单的例子从几何上阐释了交易博弈（每个卖者只能提供1个单位，每个买者只想买1个单位）的核与市场均衡的一致性。类似的结果对于任意数量的买者和卖者（他们中每一个在不同价格下可能购买或者出售多个单位）都成立，但代数推导过于复杂，因此我们这里从略。如果一些卖者或者一些买者每人都希望交易多于1单位，那么核与市场均衡的联系就要稍微弱一些。市场均衡总是位于核中，但核可能还有其他非市场均衡的配置。然而，在有大量买者和卖者、规模很大的市场中，这样的例外配置相对较少，核与市场结果事实上是一致的。由于该结果来源于相当高级的微观经济学理论，我们这里只能是稍加提及，希望这里设计的简单例子能增强你对其真实性的信心。[4]

18.3.1 市场机制的特征

1. 有效性

核与市场均衡结果之间的一致性具有直接和重要的含义。我们在18.2节看到核配置是有效的，它最大化了总的可能剩余。因此市场均衡也必须是有效的，这一事实部分解决了市场的需要。

因为有核的存在，市场不保证收入或者财富被公平地分配。一个卖者的利润取决于市场价格高于其成本多少，一个买者的剩余取决于市场价格低于他的支付意愿多少。

2. 信息与可操纵性

典型的市场交易商不知道任何卖者的保留价格或任何买者的支付意愿。他制定出一个价格，要求交易者通过买和卖的行为透露一部分信息。他们有激励做真实的回答吗？换句话说，市场机制具有可控性吗？这是本书策略视角下的一个中心问题，并与第9章讨论的信息有联系。

首先，考虑每个买者和每个卖者只希望交易1单位时的情形。为简洁起见，考虑18.1节两买者、两卖者的例子，卖者S_1的保留价格为100。他可以装作有一个较高的保留价格（通过提出不出售房产，直到市场交易商宣布价格适当高于100）或者一个较低的保留价格。也就是说，他可以假装他的供给曲线与事实上的不同。这符合他的利益吗？

正如市场交易商不知道任何人的保留价格或者支付意愿，S_1也不知道其他任何人的这一信息。因此，他不知道市场出清价格会是多少。假设他假装他的保留价格是120。市场出清价格可能超过120，这种情况下他可能决定出售，就像任何情况下那样，并且他将获得市场价格，就像他显示他的保留价格是100时他的可能所得那样；他的夸大之词可能并不会造成什么不同。但市场出清价格也可能位于100至120之间，这种情况下市场交易商可能不会采纳S_1出售的要求，S_1将可能会失去一笔有利润的交易。在这个例子中，夸大只会损害自身利益。类似

地，他可能保守地说出他的保留价格，在市场交易商报出 80 的价格时提出出售。如果市场出清价格最终低于 80，保守报价可能没有影响；但是如果价格定格于 80 到 100 之间，那么市场交易商将会按 S_1 的要求交易，从而使 S_1 遭受净损失。因而，保守和夸大一样只会损害自身利益。换句话说，真实地披露价格是 S_1 的占优战略。对所有买者和卖者都是如此。

这个推断与第 16 章中关于真实投标是私人价值第二高价密封投标拍卖的占优策略的观点相类似，这并非巧合。关于第二高价密封投标拍卖最重要的一点是，你赢得拍卖时所支付的并不取决于你的投标价，而取决于其他人的投标价。过高或过低的投标价只会对你获胜的机会产生负面影响。类似地，在市场中，如果你想出售或者购买商品，价格并非取决于你的保留价格或者支付意愿，而是取决于其他人的保留价格和支付意愿。不实地披露价格只会对你达成交易的机会产生负面影响。

如果某个人有多个单位进行交易且是能影响价格的大交易者，上述论断将不能成立。举个例子，一个卖者的保留价格可能为 100，准备出售 15 单位。价格为 100 时，可能有 15 个人愿意购买，但卖者可能假装只有 10 个单位。于是市场交易商可能相信存在短缺，他将制定更高的市场出清价格以使买者愿意购买的数量与卖者准备出售的数量相匹配。如果买者的支付意愿随可售数量的下降而提高得足够快的话（即需求曲线是陡峭的），那么该卖者以更高的价格出售 10 个单位可能比以较低的价格出售 15 个单位获得的利润更高。不实地披露信息使他具有优势，在这种情形下，市场机制是可操纵的。用经济学术语讲，这就是卖者行使**垄断力量**（monopoly power）的一个例子。通过以一个特定的价格提供低于实际拥有（或者生产）的数量，卖者抬高了市场均衡价格。类似地，多个单位的购买者可保守地披露他的需求数量以降低价格，因而在他实际购买的数量上获得更多的剩余。此时，他在行使**买方垄断力量**（monopsony power）。

然而，如果市场足够大以至于任何单个卖者企图行使垄断力量只能微小地改变价格，那么这样的行为并不能给卖者带来利润。当每个卖者和每个买者相对于市场非常小的时候，每个人的垄断力量可以忽略不计，用经济学的术语来讲，就是竞争变成完全的了；用我们博弈论的术语讲，就是市场机制变成**不可操纵的**（nonmanipulable）或者**激励相容的**（incentive compatible）了。

18.3.2 实验证据

我们已描述了市场均衡形成的过程。然而，就像前面注意到的，在这一过程中起中心作用、报价、搜集买和卖的提议、调整价格直到市场出清的市场交易商在大多数市场中并不真实存在。相反，买者和卖者必须搜寻交易伙伴，并努力与他们达成交易。这个过程有些像核的理论中联盟的形成与重新形成的过程，但更具偶然性，更加无组织。这样的过程能得到市场出清的结果吗？

通过对现实市场的观察，我们已经发现许多市场运行良好的例子，也发现了许多市场失灵的例子。但是前者与后者在许多方面情况都不一样，而且许多这样

的差异并不容易看出来。因此，这一证据无助于我们理解市场是否和如何能解决多个买者和卖者为达到均衡价格而协调彼此行为的问题。另一方面在实验室做实验允许我们在可控制的条件下观察结果，只需改变一些条件——其效应是我们所要研究的。因此实验室实验是判断市场机制作用的一种有用的方法和一项有价值的证据。

过去的半个世纪里，已进行了大量此类实验。它们的目标和方法不同，但都是将受试者集合起来，对现实的或虚拟的“物品”进行交易。每个人都被告知其自身的估价并且被给予现实的激励按照其估价行动。举个例子，完成交易的准卖者会被给予货币奖励，数额等于其在交易中获得的利润或剩余，即其交易价格减去被指定的成本的余额。

这些结果总的来说支持了理论。在市场中一方的数量只有 2 个或者 3 个的实验中，运用卖方垄断力量或者买方垄断力量的意图是可被观察到的。但另一方面，即使交易者人数相对较少（如每方 5 个或者 6 个），一般而言，交易价格仍会相当迅速地收敛到实验者根据他指派给交易者的估价计算出的市场均衡。如果有什么问题的话，就是实验的结果“太好了”。即使交易者对其他交易者的估价一无所知，不了解市场均衡的理论，先前也无交易经验，他们也会相对较快地达到均衡配置。当前的研究焦点发生了转移，已经由考察市场机制是否起作用变成质疑为什么它运行得比我们预期的更好。[5]

有一个明显的例外，即资产市场。决定购买或者出售一项长期资产，不仅需要考虑价格与你自己的估价之间的关系，还要考虑未来价格的预期变动。即使你不期望你的 dot. com 的股票会派发任何红利，你也可能购买并持有一段时间，如果你相信届时你可以以足够高的价格卖给其他人，从而实现资本获利。你可能接受这种信念，因为你观察到价格已经上涨了一段时间。但是价格上涨可能是由于其他人同样期望通过再出售获得资本利益因而也在购买 dot. com 的股票。如果许多人同时试图卖出股票以获得那些预期的收益，价格就会下跌，收益就无法实现，期望就会暴跌。于是，我们可以说 dot. com 的市场经历了投机泡沫（speculative bubble）的破裂。实验里的市场中，这样的泡沫与现实中的一样是可被观察到的。[6]

18.4 夏普利值

有效性是核的一个重要的令人满意的特征，但是它也有一些不令人满意的特征。最基本地，存在没有核的博弈，而许多其他的博弈的核非常大，它由全部的结果构成，因此这个概念不能唯一地决定一个结果。我们已构造出能在这些方面做得更好的另外一些概念。最有名的就是以加利福尼亚大学洛杉矶分校的洛依德・夏普利（Lloyd Shapley）的名字命名的夏普利值。

和第 17 章讨论的谈判的纳什解一样，夏普利值是一个合作概念。类似地，它也建立在一组公理或者假设（我们已讨论过，这些都是一个解所必须满足的条

件）上。同时，唯一解必须满足所有这些公理。特征是否令人满意是一个评判问题。其他的博弈论专家已详细说明了这些特征的其他组合，其中每一个组合都有自己的唯一解，但是不同组合的解是不同的，不是所有的特征都同等地令人满意。也许更重要的是，每个参与人并不是根据这些特征的外在的吸引力而是根据最终结果对之进行评判的。如果一个结果让一些参与人不满意，那么他们就会认为坚持另一场博弈会让他们做得更好，提议的那个解也就根本不会成立。

因此，任何合作博弈的解不应被视为博弈的可能的实际结果（仅基于那些看起来符合某些博弈论专家需要的或者抽象的特征）。我们更应该根据合作博弈解对分析和预测的实际功用来判断它的准确性。从实际的角度看，夏普利值的效果更好。这里我们不基于概念做深入的理论分析，而只是简单地阐述公式，扼要地加以解释，适当地推演，并在实例中加以说明。

假设一个博弈有 n 个参与人，其特征函数为 $v(C)$，它是联盟 C 可为其成员保证的总支付的最小值。**夏普利值**（Shapley value）就是对参与人 i 的支付 u_i 的配置，u_i定义为

$$u_i = \sum_{C} \frac{(n-k)!(k-1)!}{n!}[v(C) - v(C - \{i\})] \tag{18.1}$$

其中 $n!$ 表示 $1\times2\times3\times\cdots\times n$ 的积；上式是对所有包含参与人 i 的联盟 C 求和，每一项中的 k 为联盟 C 的规模（size）。

其思想就是每个参与人应获得相当于他为所属每一个联盟所作贡献平均值的支付，在适当的情形下，所有联盟出现的可能性是相同的。首先考虑联盟的规模 k，它的取值范围是从 1 到 n。因为所有规模的可能性相同，所以一个特定规模的联盟出现的概率为 $1/n$。那么在规模为 k 的联盟中，从剩下的（$n-1$）个参与人中选择参与人 i 的（$k-1$）个同伴，这样的选择有

$$\frac{(n-1)!}{[(n-1)-(k-1)]!\ (k-1)!} = \frac{(n-1)!}{(n-k)!\ (k-1)!}$$

种方式。上式的倒数为任何一种选择的概率。将倒数与（$1/n$）合并得到一个规模为 k 的特定联盟 C 包含成员 i 的概率，即出现在公式（18.1）右边求和符号后每一项的联合因子。与因子相乘的是联盟 C 的安全水平 $v(C)$与参与人 i 离开联盟后剩余成员的安全水平 v（$C-\{i\}$）之差。这一项测度了参与人 i 对联盟 C 的贡献。

每个参与人的支付应与其贡献相称的思想具有相当大的吸引力。最重要的是如果一个参与人必须经过努力才能产生这个贡献的话，那么这样一个支付方案会给他恰当的激励以进行努力。用经济学的术语来说，每个人对经济的增量贡献称为**边际产出**（marginal product），因此将支付与贡献联系起来的观念称为**边际生产力分配理论**（marginal productivity theory of distribution）。市场机制会根据每个参与人的贡献自动给予回报。夏普利值可理解为实现这个原理（取得与市场同样的结果）的一种方法，至少当实际市场不可能如此安排时可以近似实现，比如在一个大组织内的一些小联合体中。

这个公式比边际生产力报酬的一般性理论包含了更多内容。将一个成员为所

属的所有联盟所作的贡献取平均值，并且假定所有联盟出现的可能性是相同的，是一种非常具体的做法，现实中不一定能找到对应的例子。尽管如此，夏普利值常常产生具有一些现实特征的结果。我们现在考察夏普利值在实际中的两个特殊的例子：一个是政治的；一个是经济的。

18.4.1 立法机构和委员会的权力

假设有一个由红、蓝、绿、棕四个政党构成的100人的立法机构。红党有43个席位，蓝党有33个，绿党有16个，棕党有8个。每个政党都是一个一致的集团，它们一起投票，因此每个政党都可被视为一个参与人。任何法律的通过都需要大多数人的同意；因此，没有哪个政党可以在未得到其他政党的帮助下通过一项法律，也没有哪个政党可以独立起决定性影响。在这种情况下，一个政党的权力取决于它在多数者联盟形成过程中的作用如何关键。夏普利值正好提供了对这种权力的测度。因此在这种背景下，它也被称为**权力指数**（power index）。夏普利和马丁·舒比克（Martin Shubik）这样使用过这个指数。约翰·班兹哈夫（John Banzhaf）设计了一个相似的指数，但关于联盟形成的假设有些不同。[7]

令包括了多数者团体的联盟的支付为1（包括大联盟当然也为1），没有多数者团体的联盟的支付为0。那么在政党 i 的夏普利值的公式中，如果联盟 C 中有多数者团体，那么政党 i 对联盟 C 的贡献 $v(C)-v(C-\{i\})$ 为1，如果联盟 C 中没有多数者团体，那么政党 i 对联盟 C 的贡献为 $v(C-\{i\})=0$。当政党 i 对联盟 C 的贡献为1时，我们说政党 i 是多数者联盟 C 中的**关键成员**（pivotal member）。所有其他情况下，即如果联盟 C 中没有多数者团体或者没有政党 i，那么政党 i 对联盟 C 的贡献为0。

我们现在可以列举出所有多数者的联盟和它们的关键成员。4种由单一政党组成的可能的（尽管微不足道）“联盟”中，没有任何一个可以占多数。在6种由两个政党组成的可能的联盟中，有3种占多数，即{红党，蓝党}、{红党，绿党}、{红党，棕党} ——在任何一种可能性中，两个成员都是关键成员，因为缺少其中任何一个，都意味着失去多数的地位。4种由三个政党组成的可能的联盟中，{红党，蓝党，绿党}、{红党，蓝党，棕党}、{红党，绿党，棕党}这3种可能性都仅有红党是关键成员。第4种三个政党的联盟即{蓝党，绿党，棕党}中，全部三个政党都是关键成员。在大联盟中，没有关键成员。

夏普利值的公式中，与每个两政党联盟相对应的项的值为 $(4-2)!(2-1)!/4!=2!\times1!/4!=2/24=1/12$，与每个三政党联盟相对应的项的值也为 $(4-3)!(3-1)!=1!\times2!/4!=1/12$。

根据这一信息，我们可以计算出每个政党的夏普利值。我们有

$$u_{红党}=\frac{1}{12}\times3+\frac{1}{12}\times3=\frac{1}{2}$$

因为红党是3种两政党联盟和3种三政党联盟中的关键成员，这样的联盟中每一

个的权重皆为 1/12。类似地有

$$u_{蓝党}=\frac{1}{12}\times1+\frac{1}{12}\times1=\frac{1}{6}$$

同样地有 $u_{绿党}=u_{棕党}=1/6$，因为蓝党、绿党和棕党每个都正好是 1 种两政党联盟和 1 种三政党联盟的关键成员。

尽管蓝党成员数量差不多是绿党的两倍，绿党的人数是棕党的两倍，但是三个党的权力相同。原因是它们在形成占多数的联盟的过程中同样重要，要么其中一个党在某个时候与红党联合，要么它们三个联合起来对抗红党。

现实中，我们看到在多党制立法机构中，小政党拥有的权力的比重超过了其人数的比重，这就让我们相信夏普利值指数的强大作用。不过，夏普利值指数的确做了一些不那么有吸引力的假设，例如令所有联盟出现的可能性相同，再如令所有创建了多数者团体的政党对联盟的贡献为 1，反之则为 0。因此，我们不应该将权力指数解释为字面上的或者精确的数量测度，而只能将它解释为对政治力量水平的一种粗略表示。更加现实的分析有时候显示小政党的权力甚至比夏普利测度所暗示的更大。例如，两个最大的政党常常发生意识形态的对抗，这使它们不可能结成一个大的联盟。如果包括红党和蓝党的联盟是不可能的，那么绿党和棕党就会有更大的权力，因为它们有能力构建一个多数者团体——它们其中之一与红党结成联盟或者它们一起与蓝党结成联盟。

18.4.2 联合成本的配置

夏普利值最重要的实际应用很可能是一个联合项目其多个组成部分之间成本的配置。假设一个城镇计划建造一座多用途礼堂大楼。可能的用途是讲演厅、剧院、音乐厅和歌剧院。一座单一用途的讲演厅成本为 1 百万美元。剧院的舞台和幕后设施更复杂，成本为 4 百万美元。音乐厅比讲演厅的音响效果更好，但舞台不如剧院复杂，因此单纯的音乐厅成本为 3 百万。歌剧院同时需要舞台和音响，两项加在一起（无论是剧院加音乐厅，还是歌剧院）的成本为 6 百万美元。如果修建了歌剧院，它也能用于其他三种用途中的任何一种，而剧院或者音乐厅也能用做讲演厅。

城镇议会打算通过对使用者收费对修建成本进行补偿。每种活动的门票价格中包含了对与该活动相关的建筑物的修建成本的合理分摊，可怎样分摊呢？这是按贡献分配支付问题的一个镜像，即我们试图对产生成本的工程的最终享用者进行收费。夏普利值为这个问题提供了一个答案。对于每种用途，比如剧院，我们列举出所有可能的组合、这些组合可能产生的修建成本以及修建成本与剧院修建成本之差。这一分析告诉我们，剧院对每种组合的修建成本贡献了多少。假设包含了这种用途的所有组合的可能性是相等的，通过求平均值，我们就得到了整个成本中剧院分摊的部分。

首先，我们计算一下各种用途的每一组合的概率。剧院自身的概率为 1/4。它可以以 1/4 的概率包含在两种用途的组合中，剧院可以和其他三种用途组合，

故每种这样的组合的概率为 1/12。类似地，每种包含剧院的三种用途的组合的概率也为 1/12。此外，每种用途包含在四种用途的大联盟中的概率为 1/4。

其次，我们计算一下剧院对每种组合的修建成本的贡献。这里我们只给出一个例子。剧院和音乐厅这一组合的修建成本是 6 百万美元，而单独的音乐厅的成本是 3 百万美元；因此，6－3＝3 就是剧院对这一组合的贡献。组合与贡献的完整列举见表 18—1，其中每一行关注一种用途，最左边一列简单地标出了功能：L 代表讲演厅，T 代表剧院，C 代表音乐厅，O 代表歌剧院。接下来的四列中表示了四种用途的不同组合，每一列的上方表示了某种规模及该规模下的每一种组合方式出现的概率。这些列的每一单元格中列举了在对应列所示的规模下包含所在行所示的用途的各种组合（当某种规模下只有唯一的组合时，我们不加列举）及特定的某种用途对相应组合的贡献。最右边一列的单元格中就是以适当的概率作为比重对所有贡献求得的平均值，即夏普利值。

表 18—1　　多功能礼堂的成本配置

用途	组合				夏普利值
	规模 1 概率＝1/4	规模 2 概率＝1/12	规模 3 概率＝1/12	规模 4 概率＝1/4	
L	1	LT:4－4＝0 LC:3－3＝0 LO:6－6＝0	LTC:6－6＝0 LTO:6－6＝0 LCO:6－6＝0	6－6＝0	0.25
T	4	LT:4－1＝3 TC:6－3＝3 TO:6－6＝0	TLC:6－3＝3 TLO:6－6＝0 TCO:6－6＝0	6－6＝0	1.75
C	3	CL:3－1＝2 CT:6－4＝2 CO:6－6＝0	CLT:6－4＝2 CLO:6－6＝0 CTO:6－6＝0	6－6＝0	1.25
O	6	OL:6－1＝5 OT:6－4＝2 OC:6－3＝3	OLT:6－4＝2 OLC:6－3＝3 OTC:6－6＝0	6－6＝0	2.75

有趣的是，我们注意到虽然修建歌剧院的成本是修建讲演厅的 6 倍（“规模 1”的那一列中底部单元格中的数字与顶部单元格中的数字大小之比为 6∶1），而在组合设施中，歌剧院所分摊的成本是演讲厅所分摊的成本的 13 倍（“夏普利值”那一列中底部单元格中的数字与顶部单元格中的数字大小之比为 2.75∶0.25），这是因为演讲厅的修建并未增加其他用途建筑的修建成本，因此，讲演厅对其他所有组合的贡献为 0。另一方面，修建歌剧院增加了所有组合（同时包含剧院和音乐厅的组合除外——当然，只有一个）的成本。

我们也注意到分摊的成本的总和为 6，这是修建组合设施的总成本，也是夏

普利值的另一个有用的特征。如果你掌握了足够的数学技巧，你可以运用公式(18.1)验证所有成员的支付 u_i 加起来就是大联盟的特征函数 v。

概括一下，夏普利值通过计算每种用途的实现带来的总成本的增加来配置联合工程的成本。这样，该用途就分摊了成本的增加值——用经济学术语来表达，此即它的边际成本（marginal coast）。

这个原则也可以反过来使用。假设一些企业可以通过合作创造出比单独工作更多的总利润，它们就可以对剩余的分配进行谈判（就像我们在第 17 章所见的），但是夏普利值提供了一种有用的替代，每个企业都能获得其对联合企业的边际贡献。亚当·布兰登博格（Adam Brandenburger）和巴里·奈尔伯夫（Barry Nalebuff）在他们的专著《竞合》（*Co-opetition*）中具体应用了这种思想。[8]

18.5 公平配置机制

无论是核、市场机制或者夏普利值都不能保证支付的任何公平性。博弈理论家已经考察了其他注重公平的机制，并取得了大量的理论及实践成果。这里我们只是简单地让你们了解一下。[9]

最早的和最有名的**公平分配机制**（fair division mechanism）是“一个人分配，其他人挑选”。如果由 A 将一个蛋糕切成两块或者将现有的大量物品分成两堆，而由 B 挑选其中一块或者一堆，那么结果将是公平的。其思想就是，如果 A 的分配不均匀，那么 B 将挑选大的。知道了这一点，A 将平均分配。于是这样的行为就构成了这一序贯行动博弈的逆推均衡。

另一与此相关的机制中有一个裁判，他在蛋糕表面缓慢地移动刀子，比如从左到右。在任何一点，任何一个参与人都可以说“停”。裁判就在这一点切蛋糕，并将刀左边那块蛋糕交给说话的参与人。

我们并不要求蛋糕是均匀的（一边可能葡萄干多些，另一边可能核桃多些）也不要求一堆物品中全部物品都是同样的。参与人对蛋糕的不同部分或者不同物品的偏好可以不同，因此他们可以对各个部分的价值和谁获得更好的份额做不同的理解。于是我们可以区分两种公平的概念。首先，分配应该是均匀的，意思是 n 人中每个人都认为他的份额不少于总量的 $1/n$。其次，分配应该是**无妒忌的**（envy-free），意思是没有参与人相信其他人得到更好的份额。在只有两个参与人的情况下，两种概念是等价的。但是当存在两个以上的参与人时，无妒忌性是更强的要求，因为每人都可能认为他获得了 $1/n$ 而还有其他人获得的比他多。

多于两人时，对“一个人切，其他人选”的方式进行简单的扩展可以确保均匀性。假设有三个人，A、B 和 C。让任何一个人（比如 A）将整个蛋糕分成三块。然后问 B 和 C 他们认为哪几块是可以接受的，即大小至少是 1/3。如果两人同时拒绝的只有一块，那么他们就会把它给 A，留给 B 和 C 每人一块他们认为可以接受的。如果他们拒绝了两块，就将其中一块给 A 并重新组合余下两块。B 和 C 都必定认为剩下的这一块大小至少是原来大小的 2/3。继续用“一个人切，其

他人选”的方式在B和C之间分配这块蛋糕。于是，每人获得了他认为的至少这块的1/2，也就是最初总量的至少1/3。多人时，仍可使用这一方式，只是要进行一些数学推导。移动刀子的过程相似但更复杂，不过并不失一般性。

在多于两人时，要确保无妒忌性更加困难。这里是三人时的情况。令A将蛋糕切成三块。如果B认为其中一块比另外两块大，那么他就把这块蛋糕“削减”到至少可以为自己挑选大块蛋糕留下可能的大小。当“削减”停止后，C开始从三块蛋糕中进行挑选。如果C选择“削减”过的那块，那么B可以从其他两块中选任意一块，A获得第三块。如果C选择未“削减”过的那块，那么B一定拿走他“削减”过的那块，A得到第三块。

暂时不考虑“削减”，这种分配是无妒忌的。因为C开始从三块中挑选，他不会妒忌任何人。B至少为自己得到更大块蛋糕创造了可能，要么就从两块中拿走了较大的一块（C未拿走“削减”过的那一块），要么就从两块中挑选（C拿走了“削减”过的那一块），因此B不会妒忌任何人。最后，A认为最初的切分是均等的，因此他也不会妒忌任何人。

然后，B或者C，谁未得到“削减”过的那块，谁就将“削减”下来的分成三个部分。B或者C谁得到“削减”过的那块，谁就首先挑选，A第二个挑选，“削减”下来部分的切分者得到余下的那块。类似的理由表明整个蛋糕的分配是无妒忌的。更多人的例子和其他许多迷人的细节和应用，请读勃拉姆斯（Brams）和泰勒（Taylor）的《公平分配》（*Fair Division*，前面引过）。

如果机制缺乏别的一些必要特征而无须支付价格，那么将无法产生公平。在蛋糕或者其他物品异质、偏好多样的现实环境中，这些公平配置机制总的来说不是有效的，它们都是可操纵的。

18.6 总　　结

我们考虑一个市场，其中买方和卖方中每个人都试图与另一方中的某个人交易，与本方寻求相同交易的其他人竞争。这是一个一般性博弈的例子，其中参与人可以形成联盟，制定出联合行动的临时协议。如果一个联盟可以形成并可保证其成员的支付更高，它就会制止或者阻挡一种结果。核是由不能被任何联盟阻挡的结果构成的。

在被交换物品相同的交易博弈中，核可被解释为供需的市场均衡。当多个买者和卖者之间愿意交易的价格差异很小时，核就萎缩并决定了一个竞争性的市场均衡。这种结果是有效的。买者或者卖者很少时，会出现偏好操纵，这等同于行使卖方垄断力量或者买方垄断力量。实验室实验发现，当参与人很少，仅有五六个时，交易会迅速收敛到市场均衡。

夏普利值是一种按照参与人对可能的联盟所作贡献平均值配置支付的机制。它对联合项目成本的配置是有用的，并且是对立法机构中政党相对力量的洞见的来源。

无论核还是夏普利值都不能保证公平的结果。存在其他公平配置机制，它们常常是“一个人切，其他人选”思想的变形或者一般化。但总的来说，它们是无效的和可操纵的。

关键术语

配置（allocation）
阻挡（blocking）
特征函数（characteristic function）
联盟（coalition）
核（core）
需求曲线（demand curve）
有效配置（efficient allocation）
无妒忌分配（envy-free allocation）
公平分配机制（fair division mechanism）
可行配置（feasible allocation）
大联盟（grand coalition）
激励相容的市场机制（incentive-compatible market mechanism）
市场均衡（market equilibrium）
垄断力量（monopoly power）
买方垄断权（monopsony power）
不可操纵的市场机制（nonmanipulable market mechanism）
权力指数（power index）
安全水平（security level）
夏普利值（Shapley value）
供给曲线（supply curve）

习　题

1. 考虑 18.1 节中的两个买者 B_1、B_2 和卖者 S。卖者的保留价格是 100；B_1 的支付意愿为 300，B_2 为 200。(所有数字的单位是千美元。)
   (a) 使用谈判方法（negotiation approach）描述谈判的均衡范围和这种情况下所发生的交易。
   (b) 画出这种情况下的市场需求曲线和供给曲线，并加以说明。在你的图形中指出价格的均衡范围和交易的均衡数量。
2. 两个卖者的保留价格不同，S_1 为 100，S_2 为 150，一个买者 B 的支付意愿为 300，指出在核的配置中 x_1、x_2 和 y 的剩余分别必须满足 $x_1 \leqslant 50$，$x_2 = 0$ 和 $y \geqslant 150$。这个结果意味着当价格介于 100 和 150 时，S_1 出售房产，请加以验证。
3. 有四个卖者和三个买者。每个卖者出售 1 个单位，保留价格为 100。每个买者希望购买 1 个单位。其中一个买者的支付意愿为 400，其他两个为 300。通过画图找出市场均衡。建立所有不受阻挡（no-blocking）的不等式，求解后找出核。
4. 习题 3 中，假设四个卖者聚在一起，决定他们中只有两个进入市场，全部四人分享他们取得的所有剩余。这样做会使他们受益吗？他们中的一个或者三个进入市场会使他们获取更多利益吗？这些结果的直观解释什么？
5. 一条机场跑道计划供四种型号的飞机使用：小型团体机、窄体式商务机、宽

体式商务机和巨型机。团体机的跑道需要 2 000 英尺，窄体式需要 6 000 英尺，宽体式需要 8 000 英尺，巨型机需要 10 000 英尺。跑道的建造成本与其长度成比例。把每种型号的飞机作为一个参与人，使用夏普利值将建造 10 000 英尺跑道的成本配置给四种类型的飞机。这是配置成本的合理办法吗？

【注释】

[1] “曲线”一词一般指的是一条光滑的曲线，这条曲线是由三段直线段垂直连接而成，因此你可能会觉得奇怪。但是我们使用曲线一词来描述它，只是由于没有更好的词。

[2] 在一些欧洲国家比如瑞典的劳动力市场中，全体买者（所有者或管理方）的联盟与全体卖者的联盟（联合会）进行交易，但那不是核的概念所设想的过程的运行方式。核允许公司和工人的更小的子集形成和重新形成联盟，一般是为子集寻求更好的交易。现实中两个固定的联盟之间存在的是纯粹的谈判。

[3] A·邓肯·卢西（R. Duncan Luce）和露华德·雷法（Howard Raiffa）的《博弈和决策》(*Games and Decisions*, New York: Wiley, 1957)，特别是第 8 章和第 11 章遵循了使用可转移效用的合作博弈理论的传统处理方式。关于非转移效用的一种优秀的现代的一般性处理方式，可参见 Roger B. Myerson, *Games Theory*, Cambridge: Harvard University Press, 1991, Chapter. 9。

[4] 对核与市场均衡之间关系的一般性理论的讨论，见 Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green, Microeconomic Theory, New York: Oxford University Press, 1995, pp. 652 - 660。

[5] Douglas D. Davis and Charles A. Holt, *Experimental Economics*, Princeton, NJ: Princeton University Press, 1993, chap. 3.

[6] Shyam Sunder, "Experimental Asset Markets: A Survey," *Handbook of Experimental Economics*, ed. John H. Kagel and Alvin E. Roth, Princeton, NJ: Princeton University Press, 1995, pp. 468 - 474.

[7] 参见 Lloyd S. Shapley and Martin Shubik, "A Method for Evaluating the Distribution of Power in a Committee System," *American Political Science Review*, vol. 48, no. 3, 1954, pp. 787 - 792; John Banzhaf, "Mathematical Analysis of Voting Power and Effective Representation," *George Washington Law Review*, vol. 36, no. 4, 1968, pp. 808 - 823。

[8] Adam Brandenburger and Barry Nalebuff, *Co-opetition*, New York: Doubleday, 1996.

[9] 详细论述请参见 Steven J. Brams and Alan D. Taylor, *Fair Division: From Cake Cutting to Dispute Resolution*, New York: Cambridge University Press, 1996。

* 致谢：沈佳、马航、王伟昌、崔晋义、祝东宇、关伟雄、周宏伟、李敏、张文伟、郑平、胡平、鲁小湖、张建昌、王小芽、孙晖、罗建平、张宏宇、马民光、孙涛、顾长兵、洪云、杜伟民、王一凡、李军、崔凯、李宝民、洪勇、郑义、陶泽、胡安荣、崔剑、毛彪伟、卫民、杜兵、张小强整理了书中文献和专业词汇。

体式商务机和巨型机。团体机的跑道需要 2 000 英尺，窄体式需要 6 000 英尺，宽体式需要 8 000 英尺，巨型机需要 10 000 英尺。跑道的建造成本与其长度成比例。把每种型号的飞机作为一个参与人，使用夏普利值将建造 10 000 英尺跑道的成本配置给四种类型的飞机。这是配置成本的合理办法吗?

【注释】

[1] "曲线"一词一般指的是一条光滑的曲线，这条曲线是由三段直线段垂直连接而成，因此你可能会觉得奇怪。但是我们使用曲线一词来描述它，只是由于没有更好的词。

[2] 在一些欧洲国家比如瑞典的劳动力市场中，全体买者（所有者或管理方）的联盟与全体卖者的联盟（联合会）进行交易，但那不是核的概念所设想的过程的运行方式。核允许公司和工人的更小的子集形成和重新形成联盟，一般是为子集寻求更好的交易。现实中两个固定的联盟之间存在的是纯粹的谈判。

[3] A·邓肯·卢西（R. Duncan Luce）和露华德·雷法（Howard Raiffa）的《博弈和决策》(*Games and Decisions*, New York: Wiley, 1957)，特别是第 8 章和第 11 章遵循了使用可转移效用的合作博弈理论的传统处理方式。关于非转移效用的一种优秀的现代的一般性处理方式，可参见 Roger B. Myerson, *Games Theory*, Cambridge: Harvard University Press, 1991, Chapter. 9。

[4] 对核与市场均衡之间关系的一般性理论的讨论，见 Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green, Microeconomic Theory, New York: Oxford University Press, 1995, pp. 652 - 660。

[5] Douglas D. Davis and Charles A. Holt, *Experimental Economics*, Princeton, NJ: Princeton University Press, 1993, chap. 3.

[6] Shyam Sunder, "Experimental Asset Markets: A Survey," *Handbook of Experimental Economics*, ed. John H. Kagel and Alvin E. Roth, Princeton, NJ: Princeton University Press, 1995, pp. 468 - 474.

[7] 参见 Lloyd S. Shapley and Martin Shubik, "A Method for Evaluating the Distribution of Power in a Committee System," *American Political Science Review*, vol. 48, no. 3, 1954, pp. 787 - 792; John Banzhaf, "Mathematical Analysis of Voting Power and Effective Representation," *George Washington Law Review*, vol. 36, no. 4, 1968, pp. 808 - 823。

[8] Adam Brandenburger and Barry Nalebuff, *Co-opetition*, New York: Doubleday, 1996.

[9] 详细论述请参见 Steven J. Brams and Alan D. Taylor, *Fair Division: From Cake Cutting to Dispute Resolution*, New York: Cambridge University Press, 1996。

* 致谢：沈佳、马航、王伟昌、崔晋义、祝东宇、关伟雄、周宏伟、李敏、张文伟、郑平、胡平、鲁小湖、张建昌、王小芽、孙晖、罗建平、张宏宇、马民光、孙涛、顾长兵、洪云、杜伟民、王一凡、李军、崔凯、李宝民、洪勇、郑义、陶泽、胡安荣、崔剑、毛彪伟、卫民、杜兵、张小强整理了书中文献和专业词汇。